(2)

A FRAMEWORK FOR EVALUATING SUCCESS IN SYSTEMS ANALYSIS

Bruce F. Goeller

June 1988

P-7454

The RAND Corporation

Papers are issued by The RAND Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of RAND's contracts or grants. Views expressed in a Paper are the author's own and are not necessarily shared by RAND or its research sponsors.

The RAND Corporation, 1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

A FRAMEWORK FOR EVALUATING SUCCESS IN SYSTEMS ANALYSIS[1]

INTRODUCTION

The central purpose of systems analysis is to help parties at interest in problem situations to understand and respond effectively to those situations. In achieving this purpose, an analysis team must deal skillfully and imaginatively with many technical matters, as discussed in the *Overview* and *Craft Issues*. There are, however, some additional nontechnical matters that must be handled successfully if the analysis is to be effective, such as:

- Identifying the parties at interest in the problem situation and developing appropriate relations with them;
- Understanding the kinds of success that a study may achieve so as to help the parties at interest as much as possible.

This paper has two related purposes. The first is to appreciate and identify the roles that various parties at interest may play in a systems analysis. This discussion can help analysts in identifying the parties at interest in a particular problem situation, in developing appropriate relations with them and in communicating the study's results to them effectively.

---

[1] This paper will appear as Chapter 14 in *Handbook of Systems Analysis: Craft Issues and Procedural Choices*, edited by H.J. Miser and E.S. Quade, which is cited hereafter as *Craft Issues*. © Elsevier Science Publishing Co., Inc. A companion volume by the same editors, *Handbook of Systems Analyses: Overview of Uses, Procedures, Applications and Practice*, is cited as *Overview*.

This paper is largely an extract and partially a condensation of a draft that will appear as a RAND Note. The writing of this Note has been sponsored by the Arroyo Center, the U.S. Army's Federally Funded Research and Development Center operated by The RAND Corporation. Small portions of this paper have previously appeared in B.F. Goeller and the PAWN Team, "Planning the Netherlands' Water Resources." *Interfaces* 15(1), pp. 3-33, copyright © 1985, The Institute of Management Sciences and the Operations Research Society of America, 290 Westminister Street, Providence, Rhode Island 02903.

The second purpose is to present a framework that describes the different kinds of success that a systems analysis may achieve and the different parties' possible criteria for evaluating success. This framework for evaluating success can aid analysts in setting realistic success goals for a study and in pursuing them effectively so as to avoid pitfalls. In addition, understanding gained from the framework can aid analysts in setting up and managing studies. For example, the framework has implications for recognizing circumstances that are such harbingers of failure that they raise doubts as to whether a study should be undertaken, for anticipating and limiting the handicaps that might hamper the success of an analysis, and for staffing the study in ways that may promote the use as well as the technical quality of its results.

The next section introduces three examples that will be used to illustrate many of the points this paper treats. The third section discusses the parties at interest and the fourth section describes the framework, identifying the kinds of success and various parties' criteria for measuring them. The final section offers some conclusions concerning success in systems analysis.

ILLUSTRATIVE CASES

In order to clarify and illustrate its general points, this paper makes use of three systems analysis studies in which the author was a project leader. While these three studies do not exemplify the full range of what may be involved in systems analysis work, they are all complex and various enough to demonstrate the issues of concern in this paper.

The purpose of this section is to offer a brief introduction to each case and list where it is discussed further in this volume and elsewhere. The reader would be well advised to refresh his memory of these cases before proceeding to the rest of the paper.

The San Diego Clean Air Project (SANCAP)

For some time air pollution has been a significant problem in the United States, particularly in the air basins surrounding urban areas, which accrue pollution from industry, electric power generation, and transportation systems. In 1970 the U.S. Congress passed the Clean Air Act Amendments setting various air-quality standards for air basins in the United States and requiring that the states prepare and promulgate air-quality implementation plans for each region within their boundaries to meet the standards by 1975. The federal Environmental Protection Administration (EPA) was given major responsibility for reviewing these plans, evaluating their ability to meet the air quality standards, and penalizing states for shortfalls.

In 1970, when the amendments passed, San Diego County in Southern California violated the standard for oxidant, its most difficult species, on 226 days. In 1971 the County Office of Environmental Management concluded that, even after the new controls legislated for new cars and other sources had been implemented, the region would still be violating the standards substantially in 1975, the deadline year. Recognizing that additional controls were needed and that the overall impact on San Diego County could vary considerably with the control strategy selected, the county in late 1971 submitted a grant application to the EPA to support a thorough evaluation of alternative strategies for meeting air-quality standards in 1975 and subsequent years. Concurrently, The RAND Corporation proposed to the EPA a research design for building upon past impact assessment work to create a new methodology for evaluating urban transportation and environmental strategies. EPA then authorized a grant to the county with the understanding that the principal research activity would be subcontracted to RAND, funded by approximately two-thirds of the grant. After negotiations between the county and RAND to determine the work assignments for both organizations, a research contract with RAND for a one year long study was approved by the county in late May, 1972.

The resulting San Diego Clean Air Project (SANCAP) was undertaken to analyze alternative strategies in terms of a comprehensive set of impacts on the quality of life in San Diego and to help identify the most promising strategy for implementation. The project was a joint venture of the County Office of Environmental Management and the RAND Corporation, and involved nearly five and a half person-years of effort by RAND and about three by the county.

The county's role in the research and analysis phase of the project was to (1) manage the joint project, (2) develop much of the regional data base, (3) prepare initial sets of alternative emission-reducing tactics, and (4) analyze institutional and implementation issues.

The RAND Corporation's role was to (1) develop analysis methodology, (2) define and combine the alternative pollution-reducing tactics into cost-effective mixes (strategies), (3) evaluate quantitatively the environmental, service, economic, and distributional impacts of the strategies, and (4) develop and apply presentation techniques for the findings.

After analysis results were presented to the public and to policymakers for selection of the preferred alternative, the county's role in the implementation planning phase of the project was to be the revision of the San Diego Air Quality Implementation Plans to reflect the preferred alternative.

The alternative strategies considered in the study were mixes of *fixed-source controls*, which use technological and managerial controls on emissions from fixed sources, including not only smokestacks and organic solvent users, but also aircraft, because their significant emissions occur within airfields; *retrofit devices and inspection/maintenance policies* for vehicles; and *transportation management*, which includes bus-system improvements, carpooling incentives, and gasoline surtaxes or rationing to reduce the person-trips and vehicle miles that generate vehicle emissions.

The project's analysis results were briefed to county policymakers in May 1973, and the final report (Goeller et al., 1973a) was circulated as a draft in early July and published in December, 1973. Mood (1983)

and section 10.4 in *Craft Issues* give a fairly extended summary of the SANCAP analysis and its findings. Subsequent sections of this paper discuss what the various parties at interest did with the results and the consequences for San Diego County.

The Policy Analysis of the Oosterschelde (POLANO)

In 1953 a severe storm flooded much of the Delta region of the Netherlands, killing several thousand people and inundating 130,000 hectares of land. In 1954 the Dutch government embarked on a massive construction program for flood protection. By 1975 the new dams, dikes, and other works were nearly complete for all Delta estuaries except the largest, the Oosterschelde. There the program was interrupted by controversy, as goals other than security began to compete for attention.

The original plan had been to construct an impermeable dam across the nearly 9-kilometer wide mouth of the Oosterschelde, thereby closing it off from the sea completely, a step that would have entirely changed the rare ecology, wiped out the local oyster and mussel fishing industries, and produced many other effects. People with a special interest in protecting the fishing industry or preserving the natural environment voiced strong opposition. Those primarily concerned with safety, however, supported the original plan.

As a response to enormous controversy, in November 1974 the Dutch Cabinet directed the Rijkswaterstaat, the government agency responsible for water control and public works, to prepare a report within 18 months on the technical feasibility, financial costs, and construction time of an alternative approach: constructing a *storm-surge barrier* in the mouth of the Oosterschelde. The barrier was to be a flow-through dam with large gates that would be closed during severe storms, but that would be open in normal weather to allow a reduced tide to pass into the basin, the size of the tide being governed by the aggregate size of the opening in the barrier.

At the same time, the Cabinet, after heated debates in Parliament, also declared that the storm-surge barrier would be built--provided that the study showed that the barrier could meet a set of specified

conditions: it must be technically feasible, must be completed by 1985, must cost no more than a stipulated amount, and must provide protection against a storm so severe that it might be expected to occur only once in 4000 years. Unless these conditions were met, the original plan would supposedly be carried out. This "rather unusual political maneuver," Goemans (1986, p. 6) points out, temporarily "cleared the air but at the same time set the stage for stormy weather one and a half years later."

The Rijkswaterstaat found itself in a dilemma. Although it found the barrier feasibility study a stimulating challenge, it had previously "defended the original plan and warned that a storm-surge barrier ... might be beyond the limits of technology..." (Goemans, 1986, p. 6). Thus it feared the conclusions of its report would be greeted with suspicion, regardless of what they were. If the report concluded that all conditions could be met, there might be suspicion that the Rijkswaterstaat had been too optimistic because it was eager for the challenge of the barrier, which would produce cost and schedule overruns if it was started. On the other hand, if the report concluded that some condition(s) could not be met, there might be suspicion that the Rijkswaterstaat had been too pessimistic because it still favored the old plan.

To avoid this dilemma, the Rijkswaterstaat decided to start, in parallel with the technical feasibility study, a policy analysis study that would compare the consequences of the different approaches to the fullest possible extent. The *Policy Analysis of the Oosterschelde* (POLANO) project was established in April 1975 as a joint research venture between the RAND Corporation and the Rijkswaterstaat. The Rijkswaterstaat asked RAND to help because it had extensive experience with similar kinds of analysis (such as SANCAP) and for several years it had been working with the Rijkswaterstaat on other problems. Also, the involvement of a nonprofit corporation from the U.S. was expected to enhance the study's credibility because the U.S. was then perceived as the world leader in dealing with environmental issues.

The project began with one year of analysis, during which each organization spent about eight person-years of effort on joint research, concentrating on different but complementary tasks. RAND's primary task was to develop and then apply a methodological framework for predicting and comparing the many possible consequences of the alternatives; RAND's other tasks were to help the Rijkswaterstaat staff coordinate their various study activities on the Oosterschelde (including the feasibility study) by showing interrelations and identifying data problems, and to make them familiar with policy analysis techniques by participating in joint research. The Rijkswaterstaat's primary tasks were, on the basis of special engineering and scientific studies, to develop a specific design for each alternative approach, to analyze the consequences of the designs in which it had special expertise (such as the effects on salinity), and to provide data, as well as assistance, for the methodology being developed with RAND.

The study considered three major alternatives (each with variations within its major design concept):

- A dam across the mouth of the Oosterschelde, closing it off from the sea completely.
- A storm-surge barrier at the mouth of the Oosterschelde that would be closed during severe storms but that would be open under normal weather conditions to allow a reduced tide to flow into the estuary.
- A system of large dikes around the estuary's perimeter that would leave the estuary open in order to maintain the original tidal conditions.

There was an obvious fourth alternative--to do nothing, except possibly providing for enhanced storm prediction complemented by disciplined evacuation procedures and indemnification funds to compensate those experiencing storm losses--but such a possibility was not politically acceptable to the Dutch public, and so was not given serious consideration.

Seven categories of consequences were considered for each alternative: financial costs, ecology, fishing, shipping, recreation, national economy, and regional effects. Within each category there were several types of consequences to consider.

In April 1976, RAND presented an all-day briefing to the Rijkswaterstaat describing the methodological framework that had been developed and summarizing the results of the POLANO analysis. The Rijkswaterstaat combined this work with several special studies of its own and, in May 1976, submitted its report (Rijkswaterstaat, 1976) to the Cabinet. Later sections of this paper describe what the Cabinet and Parliament did with the report and what has happened in the Oosterschelde.

The POLANO results and approach are discussed the *Overview* (pp. 89-109). Goeller et al., (1977) provide a much more comprehensive summary, together with an overview of the extensive supporting analyses.

The Policy Analysis for Water Management of the Netherlands (PAWN)

Historically, the water management problem for the Netherlands has been too much water, not only during storms but also in normal circumstances, as much of the land lies below the level of the North Sea. But in recent decades the Dutch have faced the less dramatic, but no less urgent, problem of too little fresh water and too much pollution brought on by increased industrialization and a burgeoning population with a high standard of living.

Much of the Netherlands' wealth is derived from crops grown on irrigated land. Agriculture is by far the largest user of fresh water in the Netherlands, so water shortages can cause large economic losses.

The Rhine River, which enters the Netherlands from West Germany and flows through the country, is the Netherlands major source of surface water for agricultural irrigation and other purposes. Unfortunately, it also brings with it a substantial amount of pollution. Salinity, which has both foreign and domestic causes, is the Netherlands' most serious water quality problem, and eutrophication--heavy growths of algae in relatively stagnant waters of lakes and reservoirs--is the next most pressing and widespread.

Along with several other rivers and numerous Dutch canals, the Rhine is also a major artery for the inland shipping fleet of Western Europe on its way to Rotterdam Harbor, the world's busiest port. Low water levels in the rivers and canals can cause serious shipping delays and economic losses because only partially laden ships can navigate the inland waterways.

In 1976 a severe drought cost the Netherlands more than $2500 million in agricultural losses alone, over four percent of its gross domestic product, while worsening water quality problems. While at this time the supply of fresh surface water was adequate except in dry years, the country faced the prospect that the supply would be inadequate in normal conditions by the late 1980s. Moreover, the supply of groundwater, which was very popular because of its quality, was already inadequate.

Facing such water-management problems, and others not mentioned here, the Rijkswaterstaat--which, of course, had had a successful experience with the POLANO project--commissioned an analysis that it hoped would provide a basis for a new national water-management policy. Begun in April 1977, the *Policy Analysis for the Water Management of the Netherlands* (PAWN) project was conducted jointly by The RAND Corporation, the Rijkswaterstaat, and the Delft Hydraulics Laboratory (a leading Dutch research organization).

The PAWN project was a major undertaking. Considering both research and documentation, it directly involved over 125 person-years of effort (more than one-third in data gathering), about 48 by RAND.

PAWN's primary tasks were to:

- Develop a methodology for assessing the multiple consequences of possible water-management policies.
- Apply it to generate alternative water-management policies and to assess and compare their consequences.
- Create a Dutch capability to conduct further analyses of this kind by training Dutch analysts and by documenting and transferring the methodology to the Netherlands.

After more than two and a half years of research, the PAWN final briefing described the methodology and joint results in December 1979 to an audience of several hundred senior representatives of governmental and private organizations concerned with water resources, environmental quality, and the economy.

Although this final briefing marked the end of the analysis phase for RAND, the documentation phase continued for several years. The Rijkswaterstaat wanted unusually thorough and extensive documentation, both to support the new policy and to use in training new analysts and in performing new studies.

The Rijkswaterstaat combined results of the joint PAWN analysis with results of its own (performed with the PAWN methodology by Dutch analysts trained in PAWN) to draft its highly detailed new national policy document on water management. This document went to Parliament for formal approval in November 1984. Subsequent sections of this paper describe what was done with the PAWN results and the extent to which they affected water management in the Netherlands.

Section 6.3 of *Craft Issues* offers an overview of the work with particular attention to the issue of screening the many alternatives that could have been considered in order to arrive at a small set for detailed analysis. Goeller and the PAWN Team (1985) offer a more extensive overview; Goeller et al., (1983), a volume of almost 500 pages, describes comprehensively the methodology and results and offers an entry to the large body of literature supporting them.

The next two sections identify many of the parties at interest in these three cases, describe the different kinds of success that may be achieved by systems analysis, and discuss how successful the cases were by various criteria.

PARTIES AT INTEREST IN POLICY OR SYSTEMS ANALYSIS

The analysis of an important problem situation may be of interest to many parties, including study sponsors, analysts, policymakers, and organizations and individuals potentially affected by the situation or what is done about it. Together, these parties largely determine the success of an analysis, for their actions and interactions govern how the analysis is conducted, what is done with the findings, who is affected, and how the effects are perceived.

There is no general prescription for identifying the parties at interest for a particular study, but it is important to make their discovery an early goal. This section's purpose is to define the potential parties in terms of the roles they perform. Understanding these roles will help discover the parties at interest in a particular study and provide context for the framework described in the next section.

This section now presents and discusses two lists, one from the perspective of the problem situation and the other from the perspective of the analysis, showing the roles of potential parties at interest. A particular individual or group may play more than one role on both lists. The lists are intended to be suggestive rather than exhaustive.

Parties Related to the Problem Situation

Some parties may be affected by the problem situation or the response to it, while others may affect the decisions about what is to be done. We can describe these parties by their roles in relation to the problem situation; they include:

- Policymaker for the problem situation
- Implementor for the policy or program
- Operator of the implemented policy or program
- Decisionmaker for the problem situation
- Responsibility taker for the problem situation (Shubik, 1984, p.22)

- Staff member for any of these types of persons
- Persons affected by the problem situation
- Lobbyer
- Advisor on the decision
- Evaluator of the implemented policy or program
- Enforcer for the implemented policy or program

For some of these roles the definition is obvious, but others deserve some discussion, with examples from the illustrative cases.

Policymaker. Although the definition of decisionmaker given below is consistent with typical systems analysis definitions (see, for example, the *Overview*, p. 167), this paper's definition of policymaker, which is generally a synonym, is narrower. This is because I find it essential to distinguish between policymaking and implementation: they have innately different concerns, they occur at different stages in the attempt to improve a problem situation, and they are usually done by different organizations.[2]

In my view, policymaking is concerned primarily with selecting among options (policies) that are relatively general descriptions of *what* is to be done. Implementation, by contrast, is more concrete and specific, for it concerns *how* to do something; that is, what actions by what institutions will bring the selected policy into being. For example, if the problem situation involves improving a regional bus system, policymaking might specify that the route spacing should be two blocks rather than four blocks while implementation might specify that the local transit company, financed by a special subsidy, should run buses on First and Third Streets rather than on Second and Fourth Streets. (See the discussion by Goeller in the *Overview*, pp. 214-216.)

This paper defines a *policymaker* as an individual, group or organization that can establish or modify policies (and programs) which, through implementation, affect the problem situation. While

---

[2] Implementation and its differences from policy selection are discussed more fully in the *Overview* (Chapter 9) and in Brewer and deLeon (1983, Chapters 2 and 7-10).

policymaking authority or responsibility is concentrated within a single individual in some situations, this is rare; it is usually distributed among multiple individuals or organizations.[3]

Thus we often find multiple actors playing three distinct policymaking subroles:

- The *nominator*, who recommends a particular option or presents a short list with several promising options.
- The *selector*, who chooses the preferred option, but not necessarily from those offered by the nominator.
- The *ratifier*, who may veto, approve, or modify the selector's choice.

As examples, consider these subroles in the illustrative cases. For POLANO, the nominator was the Rijkswaterstaat, which submitted a report to the Cabinet (Rijkswaterstaat, 1976) that compared the three disparate options considered in the analysis but, by intention, did not recommend a particular alternative. The selector was the Cabinet, which decided on the storm-surge barrier alternative. The ratifier was the Dutch Parliament, who approved the Cabinet decision to proceed with the storm-surge barrier approach, but raised the issue of possible modifications to the barrier design that, in normal weather, would permit a larger flow of water into the basin, benefiting the ecology and the fishing industry. In response, the Minister of Transport and Public Works commissioned a Rijkswaterstaat study on the implications of a larger opening in the barrier, from which the Cabinet decided, and Parliament subsequently accepted, that a modest increase (to an area of 14,000 sq meters) "would be a reasonable balance between additional costs and benefits" (Goemans, 1986, pp. 9, 10).

For PAWN, the nominator was also the Rijkswaterstaat, which used the PAWN analysis (Pulles, 1985b) to draft the national policy document

---

[3] Indeed, House (1982, pp. 40-41) and others contend that the solitary policymaker/decisionmaker may be a "myth" and question whether, at the national governmental level, "the monarchical or powerful individual paradigm is operable...."

on water management (Rijkswaterstaat, 1985) that presented one coherent policy. However, the policy was shaped in part by other government ministries, provincial agencies, water boards, industrial organizations, environmental groups, and so on, which were involved in the extensive, consensus-building discussion process typical of Dutch-style democracy. The Minister of Transport and Public Works, acting as the selector, readily accepted the policy, which was near consensus, and Parliament ratified it soon after.

For SANCAP, the analysis results were the primary analytical inputs to the deliberations of the San Diego County Task Force on EPA regulations, which served as the nominator: The task force report recommended a particular strategy, developed in the analysis, as an alternative to an EPA plan. The San Diego County Board of Supervisors endorsed the report and selected the recommended strategy to be presented at EPA hearings. The Region IX office of the EPA, acting as the ratifier, subsequently promulgated the Final Implementation Plan for the San Diego Air Quality Control Region (October 15, 1973), a plan that was similar in many respects to the San Diego strategy, but had some major differences in the transportation management controls (Goeller et al., 1973a, pp. 71, 94).[4] Subsequent revisions and compromises occurred that will not be discussed here.

Implementor. The *implementor* attempts to execute the policy chosen by the policymaker. In this process, decisions must be made and actions must be taken to flesh out the relatively general description of the selected policy; concrete details must be added and open issues resolved. Sometimes the result is consistent with the policymaker's intent, but often it differs substantially. Some differences result from adapting the policy to make it more practical or to accommodate new political concerns. Others occur because the policymakers and the implementors are usually different--different individuals and organizations, each with different values, perceptions, and incentives.

---

[4] Note that the 1970 Amendments to the Clean Air Act required the EPA to determine the efficacy of implementation plans submitted by the states. If the State of California had submitted a plan acceptable to the EPA in time to meet its various deadlines, the State would have been the selector (and the county the nominator).

Brewer and deLeon (1983, p. 20) suggest these activities as characteristic of implementation:

- Developing rules, regulations, and guidelines to carry out [policy] decision
- Modifying [policy] decision to reflect operational constraints, including incentives and resources
- Translating [policy] decision into operational terms
- Setting up program goals and standards, including a schedule of operations

To these we may add:

- Developing a complete and detailed design for the policy
- Preparing the implementation plan
- Constructing or procuring needed facilities and equipment
- Hiring and training needed personnel
- Drafting and letting contracts for other organizations to perform activities such as the previous two
- Administering contracts
- Establishing organizations or associations of existing organizations
- Operating the implemented policy or program

Of course the particular mix of activities will depend on the nature of the policy and the problem situation.

Implementation often finds different actors playing three distinct subroles:

- The *installer*, who creates the facilities or assembles the resources necessary to get the (new or modified) policy into operation.

- The *operator*, who is responsible for the day-to-day operation of the implemented policy.
- The *implementation manager,* who develops guidelines, prepares plans, administers contracts, and so on, and may select and manage the installers and operators or perform some of their functions.

For POLANO, the contractors who built the barrier and related facilities such as compartmentation dams and locks can be considered installers. But the Rijkswaterstaat was the implementation manager: Its staff selected the contractors from several competing groups; it drafted, negotiated and administered the contracts under which the contractors and various research institutes conducted their work; it established the schedule and the implementation plan; and it worked with the contractors to develop the detailed design for the barrier, which underwent several revisions while adapting to practical problems.

The Rijkswaterstaat also had several open issues to confront during implementation, which it did by initiating several smaller policy studies with the POLANO approach and then taking appropriate actions. These open issues included deciding on the exact location of the compartmentation dams and determining the type and dimensions for the shipping locks to go in the dams.

However, the largest managerial decision, and supporting study, considered an issue of operational policy rather than installation: selecting the control strategy for operating the barrier's many large gates. Such a strategy includes (1) the actions that govern the times and rates at which gates close and open, (2) the rules underlying the decisions for these actions, and (3) gathering and processing the required information. In choosing a control strategy, the familiar POLANO conflict between security and ecology was confronted again: From the standpoint of flood security the barrier should be closed as long as necessary, while from the standpoint of ecology it should be closed as briefly as possible. And the choice was complicated by uncertainties in forecasting storm behavior and resulting water levels. The strategy

choice was made largely on the basis of analysis by the BARCON (barrier control) project, a joint effort of the Rijkswaterstaat and The RAND Corporation (Catlett, et al., 1979).[5]

For the national water management policy developed through PAWN, the Rijkswaterstaat is probably the major implementation manager and operator. Among other things, it is responsible for constructing the Brielse Meer pipeline, for adopting a more stringent thermal standard to limit heat discharge into canals, and for introducing and operating a new flushing policy to reduce the salinity of the Markermeer (an enormous freshwater reservoir three-fourths the size of San Francisco Bay). But the provinces and water boards (local water management authorities, similar to school boards) are also important implementors. Under the regional approach to eutrophication recommended by PAWN, provinces are using improved versions of the PAWN eutrophication models to develop tailored combinations of control tactics for their lakes and reservoirs (Goeller and the PAWN Team, 1985, pp. 30, 31). Moreover, about half the provinces and numerous water boards have cut back their plans for new facilities on the basis of PAWN recommendations (Pulles, 1987).

For SANCAP, fragmentation of power complicates implementation, particularly if the county tries to implement its preferred strategy. Goeller et al., (1973a, p.73) note:

> When one examines the various tactics for improving air quality, he discovers how many different organizations are potentially involved in implementation. The local APCD [Air Pollution Control District] and the Navy may influence fixed source control. County government may indicate the need for a particular retrofit strategy, but only the state and Federal governments have the power to require installation. The San Diego Transit Corporation and the city of Oceanside manage the bus service in the San Diego region, but the various penalties against the automobile that might make the bus more attractive

[5] Another important operational policy decision during implementation involved developing an ecosystem management plan for each water basin being created by the construction. This required extensive negotiation between the different levels of government (national, provincial, and municipal) because the administrative structure did not correspond with the ecosystem boundaries.

are beyond their powers; additional parking fees require action by each of the thirteen cities plus the county, while additional gasoline charges require state or Federal action. And, of course, there are helpful actions by citizen groups and individuals.

Operator. On the one hand, operators are considered independent actors who play a major role after implementation is complete; on the other, they should often play a subrole during implementation, as mentioned above. The *operator* is responsible for managing and maintaining the implemented policy, whatever form it finally takes, in day-to-day operation. If the operators are different organizations or individuals than the implementation managers or installers, the result may differ from what the managers intended. The difference will depend on the adequacy of the operational guidelines (their concreteness and comprehensiveness), the character of the operators, and the nature of their incentives. It will also depend on the extent to which the operators concerns and constraints are considered during implementation planning and installation.

For POLANO, the operator is the Rijkswaterstaat Directorate for the Province of Zeeland, the location of the storm-surge barrier. In an artful attempt to minimize differences between implementation and operation, as the barrier neared completion the man who had headed the Rijkswaterstaat organization primarily responsible for constructing the barrier (the Delta Service) was appointed to head the organization that would be responsible for operating it.

For PAWN, the major implementors--the Rijkswaterstaat, the provinces, and the water boards--are operators as well as managers. And the same is true for most of the implementors in SANCAP.

But many problem situations include operators not commonly considered to be implementors. This generally occurs when the policy being implemented by one organization employs pricing measures (charges and surtaxes) or regulation measures (administrative and legal restrictions) to influence the behavior of others. Consider some examples from SANCAP. When the county and the various cities implement parking surcharges, or when the state or Federal government implement

gasoline surtaxes, they are trying to influence automobile operators to drive less and, in the long run, to buy automobiles that are simultaneously more fuel efficient and less polluting. When the Air Pollution Control District sets quotas on the permissible emissions by individual firms, it is trying to influence the operators of the firms; in the short run operators may reduce the production of goods to reduce associated emissions, while in the long run they may substitute less polluting equipment or production processes.

Price and regulation measures usually afford the affected operator considerable flexibility in choosing his response. But monitoring and enforcement activities are often necessary to insure that the operator's response has the desired effects.

Decisionmaker. In this paper's usage a *decisionmaker* may be either a policymaker or an implementor. This is equivalent to typical system analysis definitions, as mentioned before (although not all definitions seem to include the operator subrole).

Persons Affected by the Problem Situation. This role includes persons directly affected; for example, those threated by floods and those employed constructing the barrier in POLANO, and those using the water supply in PAWN. But the role also includes those indirectly affected--in future as well as present generations. Future generations may suffer deaths or deformities from the present generation's discharges of toxic metals into the water. And expenditures on the barrier or water-management facilities create jobs not only in the industries directly involved in construction but also indirectly in other interrelated industries, such as steel and concrete production.

Lobbyer. A *lobbyer* is an individual or group that seeks to influence policymakers toward a particular viewpoint. Lobbyers may be persons affected by the problem situation, persons in organizations affected by the problem situation but not directly involved in it themselves, or concerned persons outside the problem situation.

In POLANO, for example, formal and informal lobbyers included the Provincial Government of Zeeland, which wanted primarily to assure adequate flood protection for its residents; oyster and mussel fishing industry groups that wanted to maintain their livelihood; environmental

groups that wanted to preserve the Oosterschelde's ecology; and Comite Samenwerking Oosterschelde (SOS), a consortium of various groups that wanted to keep the mouth of the Oosterschelde open. (Even several years after Parliament ratified the decision to build a storm-surge barrier in the mouth of the Oosterschelde, SOS was actively criticizing the decision and lobbying for an open Oosterschelde.)

In pursuing their goals, lobbyers may apply political pressure, present arguments, provide information, or offer political support on other issues. Many are ardent volunteers rather than professional advocates paid to represent a viewpoint (that is, professional lobbyists).

Advisor on the Decision. When a decisionmaker consults an advisor about the problem situation, he is seeking help. The *advisor* may supply information, conduct analysis, recommend action, suggest political strategy, or provide emotional support.[6] It is in this advisory role that policy analysts usually appear in relation to a pending decision.

The main difference between analysts and other advisors is the degree to which their recommendations are based on problem-specific analysis, their assumptions are made explicit, and their personal preferences are set aside. (If an analyst allows his preferences to shape materially the analysis or recommendations, he is functioning instead as a different kind of advisor--or even a lobbyer--and should make this role shift, and its implications, clear to the clients, as discussed in the *Overview*, pp. 320-325.)

There are many kinds of advisors. They may be inside a decisionmaker's organization or outside. They may be private consultants or academic researchers who are considered experts on the problem situation, or staff members who are considered nonexperts. And advisors may be individuals or groups, such as a council, committee, commission, or task force.

---

[6] Goldhamer (1978) discusses the role of the advisor from the perspective of advising the leader of a nation. Benveniste (1977, p. x) discusses "how experts influence public...policy," asserting that experts such as policy analysts fill a "new social role combining political and technical dimensions."

For example, the POLANO analysts were preceded by several other advisors. According to Leemans (1986, p. 50), a study group of the Technical University of Delft developed the initial idea of a tidal dam, the forerunner of the storm-surge barrier. A special advisory committee to the Dutch Cabinet, the Oosterschelde Committee, considered hundreds of different plans for protecting the Oosterschelde and recommended an approach (Goeller et al., 1977, p. 15; Goemans, 1986, p. 5). Then an important standing committee of the Dutch government, the Spatial Planning Committee, which included top officials of most ministries, was responsible for evaluating these recommendations and reaching its own conclusions (Leemans, 1986, p. 51).

As another example, the SANCAP analysis results were used by another advisor, the San Diego County Task Force on EPA regulations, as the source of the strategy recommended to the County Board of Supervisors. And Dutch water management policymaking received advice not only from the PAWN study, but also from the ICWA, the Interdepartmental Committee for Water Management, which included top-level representatives of all government ministries concerned with water policy.

Evaluator. After the implementation of a policy or program, an *evaluator* compares the actual effects on the problem situation with the expected effects, using "established criteria".[7] An evaluator may belong to an organization involved in the policymaking, implementation or operation, or be a totally distinct and autonomous entity.

Enforcer. An *enforcer* is an evaluator with teeth; that is, when an enforcer observes an implemented policy performing unacceptably, he has power to induce policy changes. For POLANO, the Delta Institute at Yerseke is enforcer for the ecology; it can induce changes in the barrier control strategy if it adversely affects the ecology. For PAWN, different parts of the Rijkswaterstaat, under Parliamentary oversight,

---

[7]Brewer and deLeon (1983. pp. 19, 20) pose the comparison in terms of performance levels. I feel this is too narrow, for a policy or program may have major effects--including some that are indirect, some that are spillovers, and some that are unanticipated--far beyond its performance levels.

monitor and enforce compliance with water quality standards. For SANCAP, the EPA monitors regional compliance with air quality standards and control plans, and can impose sanctions that withhold federal funds or restrict regional growth.

Parties Related to the Analysis

Thus far this section has discussed parties at interest from the perspective of the problem situation. Now it takes the perspective of the system analysis: Some parties may shape the analysis while others may affect how it is perceived or used. The roles of the parties at interest in relation to the systems analysis include:

- Problem poser *(Overview*, p. 167)
- Sponsor
- Client
- User
- Member of the staff of any of these four types of persons
- Systems analysis team
- Systems analysis peer group
- Research program director
- Advisor on the analysis
- Formal reviewer
- Implementation planner

The categories where the definition is not obvious are discussed below.

Sponsor. This paper refers to the individual who commissions the work and sees to its support as the *sponsor*. For POLANO this was H. Engel, Chief Engineer-Director of the Rijkswaterstaat's Delta Service, and his predecessor, H. A. Ferguson. For PAWN it was H. M. Oudshoorn, Chief-Engineer Director of the Rijkswaterstaat's Directorate of Water Management and Water Movement. And for SANCAP it was L. Edwin Coate, Director of the Office of Environmental Management during most of the project, and his successor David Nielson, who served as Acting Director.[8]

---

[8] Stanley Greenfield, EPA Assistant Administrator for Research and

The sponsor plays a major role in shaping the analysis. He commonly selects the research organization to perform the analysis, determines the funding level, and may influence the staffing. He also participates in problem definition and reviews progress and findings.

Note that a study sponsor is not necessarily a policymaker for the problem situation. In POLANO and PAWN, the sponsoring organization (the Rijkswaterstaat) played the nominator's role in policymaking, but did not select the policy. In SANCAP, the study sponsor was not a policymaker, although the study results were an essential element in the policymaking process.

Client. Much of the literature refers to the client as meaning various things, including the sponsor. This paper uses the term *client* to mean a potential user of the study findings, such as a decisionmaker, his staff, or an interest group trying to influence the decisionmaking.

Through their actions, clients determine directly how an analysis is used. But they may also influence indirectly how the analysis is performed, as analysts emphasize issues they know or expect to be important to clients.

Client examples appear earlier in this section, as parties related to the problem situation.

Research Program Director. A policy or systems analysis is typically organized as a project, with a small research team managed by a project leader. When multiple analysis projects are under way in an organization, similar projects are often organized as a research program, managed by a program director. The program director usually selects a project leader, helps negotiate the research agenda with the sponsor, and reviews progress and findings. Based on his strategic plan for the program, the program manager may allocate resources among competing projects and determine which prospective projects will start. The strategic plan reflects such goals as building intellectual capital

---

Development, was the godfather, if not the sponsor, of the project. He realized the potentialities of a San Diego County/RAND collaboration, helped frame the project objectives and guide its evolution, and provided the EPA grant to San Diego County that funded most of the project's activities.

for the analysts, promoting synergy among projects, developing a center of expertise, and creating a foundation for growth into new areas.

Advisor on the Analysis. As it designs and conducts a systems analysis, the analysis team may receive guidance from selected advisors. Although such advisors may be solitary individuals, in complex problems they are often formal advisory groups, either study advisory groups--sometimes called steering committees--or technical advisory groups.

Study advisory groups provide guidance on the scope and emphasis of the research, that is, the issues to be considered and the relative importance of each. Because of its complexity and size, PAWN benefited from three such advisory groups. One was the Rijkswaterstaat advisory group, which consisted of senior managers from various Rijkswaterstaat Departments concerned with different aspects of water management (such as shipping or water quality). Another was the IWW, a special working group of the ICWA (the Interdepartmental Committee for Water Management) established specially for discussions about PAWN and the national policy document on water management. The third was the RAND advisory group, which consisted of a senior RAND expert on hydrology who was born and educated in the Netherlands (J. J. Leenderstse), the head of the System Sciences Department (G. H. Fisher), and RAND's Senior Vice President (G. H. Shubert). (Goeller, et al., 1983, pp. xxviii, xxix.)

By contrast, technical advisory groups provide data and expertise on the more concrete and specialized aspects of the problem. They may suggest options for solving the policy problem, identify potential pitfalls for the analysis, propose analytic strategies, and critique methodology and recommend improvements. In PAWN there were technical advisory groups on pollution, nature, shipping, and technical and managerial options for improving water management (Goeller, et al., 1983, p. xxviii). There were also individual technical advisors for particular topics in various study areas. For example, a senior researcher of the Ministry of Agriculture and Fisheries (H. Ton) spent a week at RAND reviewing PAWN's agricultural "research plan and preliminary versions of several ... models" (Petruschell et al., 1982, p. xiii).

Implementation Planners. Those who perform analysis specifically to assist implementation planning or decisionmaking are referred to here as *implementation planners*, a term I believe reflects the detailed and prescriptive nature of such activity better than its synonym "implementation analysts".

Depending on circumstances, the group of implementation planners may or may not include analysts from the team that supported the policymaking process. Such inclusion is certainly desirable, for the analysis team understands the problem situation, has experience assessing the implementability of proposed solutions, and should be well equipped to devise ways to resolve implementation difficulties. Only in this way can the policy analysis team influence implementation directly. Otherwise its influence on implementation, if any, will be indirect, arising through its influence on policy selection or the subsequent use of the study results by implementation planners and decisionmakers.

KINDS OF SUCCESS

For policy and systems analysis studies, "success" is difficult to define and measure, partly because different parties at interest have different goals and perspectives. Definitions and criteria for success exist, but they all have weaknesses as well as strengths, and some, if used alone, might mask serious defects or ignore important contributions. This section describes various definitions and attempts to synthesize them into a reasonably coherent framework for describing success in systems analysis.

Three general kinds of success can be distinguished:

- *Analytic success* considers how the study was performed and presented.
- *Utilization success* considers how the study was used in policymaking or implementation.
- *Outcome success* considers what happened to the problem situation (and those affected by it) as a consequence of the study.

On a *direct* basis, one assesses these three kinds of success with respect to the decision and the problem for which the analysis was commissioned. On an *indirect* basis, one assesses them with respect to other decisions and situations.

There is a hierarchy of dependence and difficulty among the kinds of success. Analytic success is the easiest to achieve, although it requires careful work to produce a high quality study, one that is valid, credible, and pertinent to policymaking. Analytic success is also the foundation for the other kinds of success. For utilization success, study results must not only be used in the policymaking process but should also be of high quality; utilization of poor quality results is no credit to a study. Outcome success is the most difficult to achieve, for it requires establishing that the study results were an important factor in policy selection and that implementation of the selected policy helped the problem situation.

Each kind of success clearly pertains to a different domain; the first pertains to the study itself, the second to the policymaking or implementation process, and the third to the problem situation. And each kind may apply to a different time period; for example, outcome success, if it occurs, may not be discernible until long after the analysis is completed.

The success of a study may appear different from the perspectives of the various parties at interest, for they often have different objectives. One interest group may use the results of the study in the policymaking process, while another may not. One interest group may find a particular implementation outcome to be beneficial while another considers it harmful.

With these general comments as background, this section now examines, first on a direct basis and then on an indirect basis, the kinds of success and the criteria for measuring each. The criteria are incommensurable, and their relative importance varies with the analysis and problem situation. Nevertheless, oversimplifying somewhat, the more criteria by which an analysis is successful, the greater its success.

Analytic Success

There are two general ways to determine the extent to which a study is an analytical success: formal quality control and approval of selected parties.

Formal quality control. Formal quality control attempts to apply a clearly defined process and explicit standards to evaluate the quality of a study and its findings. In my view, study quality has these main components:

- *Technical validity:* The study methodology appears sound, the analysis considers the relevant policy alternatives and the important uncertainties, and the findings follow explicitly from the analysis. (This is discussed in Chapters 13 and 15 of *Craft Issues*.)
- *Persuasive validity:* The analysis and its findings are presented cogently and clearly, without exaggeration or oversimplification. (See Chapters 13 and 15 of *Craft Issues*.)
- *Availability:* The study content is readily available, through reports or briefings, to its clients and reviewers.
- *Credibility:* The study content appears believable or trustworthy to the study clients.
- *Timeliness:* The study content is available in sufficient time to be digested and used for policymaking.
- *Pertinence:* The study content applies to clients' substantive concerns and spheres of responsibility.[9]
- *Usefulness:* The study content is potentially valuable to its clients; for example, the study provides a helpful new perspective or devises feasible ways to improve the problem situation. (Chapter 15 of *Craft Issues* considers the "worth" of a study as a composite of its pertinence and usefulness.)

---

[9] I refer to this component as "pertinence" rather than "relevance" because the latter term can have a variety of meanings in systems and policy analysis. Lynn (1978b, pp. 18-19) provides an excellent discussion of these meanings.

Unfortunately, the field of policy analysis lacks the agreed-upon standards and evaluation procedures necessary for formal quality control to be applied and accepted. (Chapter 15 of *Craft Issues* discusses quality control issues in detail.) Thus, until until more progress is made toward formal quality control, there must still be heavy reliance on the approval of selected parties as the means of gauging analytical success.

Approval of Selected Parties. Approval is innately subjective. So one who contemplates using the approval of another party to measure a study's analytical success should consider the party's perspective and possible biases, as well as the specific context in which the judgement was made. Consider these different marks of analytical success:

1. The first mark of analytic success is that the work and its findings satisfy the sponsor. SANCAP, POLANO, and PAWN all satisfied their sponsors. It seems obvious that any analysis team should want a satisfied sponsor, but closer examination suggests that the matter is not so clear. It must, unfortunately, be admitted that there has been work aimed frankly at confirming the sponsor's prior prejudices; if such work ignores relevant disturbing facts and bends the analysis to this motivation, then it clearly is bad analysis, and the satisfied sponsor is not a mark of success. It is far commoner for a meticulous and searching analysis to produce findings that disturb the sponsor, at least as they first reach him; here the unsatisfied sponsor would seem to be the mark of greater success. Of course, there are discriminating sponsors who realize that the news emerging from a searching examination of a problem situation is quite unlikely to be all good, and who respond to it by examining its basis; if they find that the basis is solid and the interpretations emerge from sound analysis, then they may be unhappy about the troubling findings but satisfied with the analysis--and thankful that the analysis has brought the troubling matters to light. In sum, for the criterion of a satisfied sponsor to be judged a mark of success demands that the context from which the satisfaction emerged be known in some detail.

2. Another mark of analytic success is that the work and its findings satisfy the analysts. Here again knowledge of the context conditions how such a mark should be viewed. On the one hand, a study that gets only a few rudimentary steps beyond the information-gathering stage may turn up an insight that causes the sponsor to view his central question in an entirely new light, enabling him to approach its resolution from a new direction; yet the analysts may remain somewhat unsatisfied with such a rudimentary analysis, even though its findings helped the sponsor. Similarly, since few systems analyses escape some sort of time constraint on their completion, a quite competent study may be forced by this limitation to draw to a close earlier than a meticulous analysis team might wish. On the other hand, an analysis may be so complete and presented in such detail as to swamp the sponsor with information, thus perhaps satisfying the analysts but leaving the sponsor less so. Too, if the drive for analytic completeness and satisfaction overrides the time constraint so far as to yield findings after the related decisions have been made, the analysts may be satisfied with the technical quality of their work but have a disappointed sponsor. I believe, however, that competent analysts should not be fully satisfied with their work if the findings are late or presented ineffectively, as in these two examples. Thus, satisfied analysts are a mark of success in systems analysis only if the surrounding circumstances support this judgment and the analysts are competent.

3. Yet another mark of analytic success is that the project satisfies the research program director. One might suppose a program director would be satisfied if a project was high quality and pleased the sponsor. But matters are not always so simple. The program director may be dissatisfied with such a project if it has a high opportunity cost; if, for example, its findings antagonize another sponsor, or its resource consumption (money or staff) delays or undermines a more important project. And he may be disappointed if the project evolves in directions that produce little of the synergy with other projects envisaged in his strategic plan.

A research program director judges the success of a project in the context of his portfolio of projects and in terms of its contribution to his program's overall goals. As an example, consider a project that failed to satisfy the sponsor because serious conceptual or methodological difficulties kept it from producing any findings. Yet the program director might consider the project successful if it led to a follow-on project for the same sponsor, built intellectual capital for other projects or identified a promising topic for future research.

SANCAP not only pleased its sponsors and reviewers, it was also particularly pleasing to its program director. It led directly to two other studies--studies of air pollution and transportation management strategies for Los Angeles--and indirectly to POLANO. POLANO led directly to both BARCON and PAWN.

4. A systems analysis study may be considered an analytic success if peer reviewers judge it to have high quality (as discussed, for example, in *Craft Issues'* Chapter 15) and if the basis of the review is appropriate to the study. Although peer review is universally acknowledged to be an essential component in judgments of quality for science and technology, the process is not without its faults, as a number of critics have made clear (see, for example, Armstrong, 1982). To detail possible faults in the peer review process is unnecessary here, but several of its potential limitations for systems analysis studies should be discussed.

One limitation concerns novelty. A reviewer necessarily uses his knowledge of the craft of his subject as the background for his review. Thus, if the study being reviewed falls within this envelope of craft knowledge, then the existing standards of quality with which he is familiar apply, and are an appropriate basis for judgment. If the work is novel in important respects, however, the existing criteria may not be appropriate, and, if used, will inevitably find the work to be below standard in some ways. It follows that the peer review system tends to be conservative in its effects, unless the reviewers can generate and apply the new criteria of quality that pioneering work may demand. (See Ravetz, 1971, pp. 182-184.) Since an important systems analysis almost

inevitably plows new ground in its analysis, it is especially important for its review to be based on criteria that are appropriate to the work and its relation to its problem-situation context. It is unfortunately the case that reviewers do not always adhere to this principle; the demand for rigor rather than relevance has even been observed (*Craft Issues*, Section 12.4).

Another limitation concerns multidisciplinarity. A systems analysis commonly involves multiple disciplines. As the number of disciplines (and perhaps the synthesis among disciplines) increases, it becomes more difficult to obtain adequate peer review. Finding individual reviewers who are simultaneously expert in systems analysis and all relevant disciplines is next to impossible. And using multiple reviewers with complementary expertise is often insufficient: An impractical number of reviews may be required to cover all the disciplines involved, and the reviews may well be piecemeal judgments rather than comprehensive and holistic appraisals.

Perhaps the most serious limitation concerns competence. Because systems analysis is more art than science, it can be difficult to get reviewers with the special competence required. The reviewers must not only be fairminded and unbiased, but their expertise in the disciplines involved and in the craft of systems analysis must also be equal, and preferably superior, to that of the analysis team; otherwise one risks obtaining reviews of questionable worth, as when a Salieri critiques a Mozart or a Warhol critiques a Picasso. The reviewers also must be sufficiently knowledgeable about the problem situation that they can properly judge a study's policy findings as well as its methodological basis.

The methodology and results of PAWN and SANCAP were approved by the appropriate technical advisory groups, and the RAND reports on all three studies received a careful technical examination from at least one reviewer unaffiliated with the project.

PAWN provides two unusual illustrations of approval by reviewers. First, in a 1984 international prize competition, PAWN won The Institute of Management Sciences' Award for Management Science Achievement (the Edelman Prize).

Second, the Minister of the Netherlands Ministry of Transport and Public Works, Mrs. N. Smit-Kroes (1984) made the following statement about PAWN:

> [T]he various cost and benefit numbers have been carefully examined without challenge by various other government ministries, provincial and local government agencies, private industries and associations, and public interest groups.
>
> . . . On the basis of extensive examination [the PAWN methodology] appears to give credible results and to represent a substantial advance in the state of the art.

Despite the potential limitations discussed above, peer review is clearly valuable for judging a study's technical validity. And peer reviewers can usually detect when statements of study findings are so unclear or exaggerated as to undermine persuasive validity. But they are not well equipped to predict whether the various study clients will deem the statements credible or will consider the findings pertinent and useful. For these components of quality I believe the study's clients are generally much better judges than the analysts' peers.

5. Another mark of analytic success is that the study satisfies its clients. But the meaning of this mark is blurred by client multiplicity. A study's multiple clients may reach different judgments as to analytical success because of their different goals and perspectives. Of the main components of study quality, usefulness probably produces the most disagreement. Clients who agree that a study is valid (perhaps on the basis of the same peer review), credible and pertinent may nevertheless disagree on whether the study is useful and thus successful. Client A may consider the study successful because it devises a new solution to the problem that he finds attractive, while client B may do so because its findings support a solution he already favors, as was the case with SOS and POLANO.

Goemans (1986, p.8) notes the following client reaction to POLANO: "Although nobody [except the sponsor] had asked for POLANO, there was widespread appreciation for the [Rijkswaterstaat (1976)] report which showed the issue in its broader context."

There are even instances where clients' biases dominate their judgments of analytical success. When valid study findings, through distortion or presentation out of context, can be made to appear to support his position, a client may consider the study successful. And a client may consider a study successful if its findings support his position, even though they are of dubious validity.

Utilization Success

While analytic success concerns how the study was performed and presented, utilization success concerns how it was used. For utilization success, the study must be *used* in the policymaking process.[10] Otherwise, its influence can only be academic and its monuments merely journal citations and reports gathering dust on bookshelves.

Various authors conclude, on the basis of empirical investigation, that policy research utilization has been seriously underestimated because "use" is too narrowly defined.[11] Numerous researchers have defined and refined concepts of knowledge use in policymaking.[12] I have synthesized several ideas (some from the literature and some of own) into a framework for characterizing a policy research study's utilization. This framework addresses three component questions: namely: (1) who uses (2) which elements of a study (3) for what purpose? I now describe the framework by addressing each question in turn.

---

[10] Recall that the study must also be high quality (an analytic success), as utilization of poor quality results is no credit to a study.

[11] For example, Pelz (1978, p.346), Weiss and Bucuvales (1980, pp. 10-11), Larsen (1985, p. 145), and Whiteman (1985b, p.204).

[12] They include Rich (1975, 1977, 1981), Caplan et al. (1975), Weiss (1977a, 1977b, 1980), Pelz (1978), Larsen (1980, 1985), Dunn (1983), Beyer and Trice (1983), and Whiteman (1985a, 1985b).

Study Users. A policy study cannot be considered a utilization success if it is used only by users who do not participate in the policymaking process (such as students and research managers). Rather, for a study to be a utilization success, at least one of its users must participate in policymaking. Such users may be policymakers (and sometimes implementors), their staff, advisors on the decision, or lobbyers.

One might contend that a study's utilization success increases with the number of users. But matters are not that simple. Some users are more important than others; for example, use by one policymaker probably has greater significance than use by several staff members or lobbyers. Also, collective (group) users are generally more important than individual users, even when the same people are involved. For example, a study would probably be considered a greater utilization success if it led a group of policymakers to a common position than if it led the individual members to divergent positions. The distinction between individual and collective users focuses on mutuality of use and shaping of consensus, whether for problem formulation or policy selection. (Dunn, 1983, distinguishes individual and collective use, but with a somewhat different slant.)

Elements Used. A policy study typically contains many elements of knowledge that might be used in the policymaking process. The types of elements include terminology, issues, arguments, concepts, models, alternatives, estimates, findings, and recommendations.

A user may elect to use some study elements and disregard others. Usually elements are used or disregarded on the basis of whether they appear pertinent and useful. Sometimes, however, elements that appear useful are deliberately disregarded. This may happen because they raise political difficulties, as Larsen (1980, p. 430) suggests, because they lack credibility, or because they are less compelling than knowledge from other sources.

Mere consideration of an element of knowledge in the policymaking process does not constitute use. Rather, the element must affect a user's thinking about the problem and its solution (conceptual use) [13] or his observable behavior toward them (instrumental use).[14]

---

[13] Of course the inclusion of effects on thinking greatly complicates the empirical measurement of knowledge use. (One cannot

One might contend that a study's utilization success increases with the number of elements used in policymaking. But, as with users, some elements are more important than others; the use of a recommendation, for example, is probably more significant than the use of particular terminology. In principle, a study could be considered a utilization success if only one individual element is used. In practice, more than this would usually be needed, unless the element is quite important. However, given the usual number and variety of elements, it is unrealistic to expect that every element in a study would be used.

It is also unrealistic to expect that a particular element would be used exactly as it was presented in the study. Adaptation occurs naturally during the utilization process. Users, sometimes with assistance from the analysts, often modify study elements to fit their needs (Berman and McLaughlin, 1977; Larsen, 1980, p. 428). Ideally, utilization is an interactive process that involves analysts as well as users.

Purpose of Use. Knowledge from a systems or policy analysis may be used for various purposes. Each purpose represents a different kind of utilization in the policymaking process. A particular study thus may be a utilization success for some purposes and a failure for others. We

---

accurately measure such effects by counting citations of a study or by administering a simple questionnaire.) However, I shall not discuss this topic further here because my focus is not on measuring knowledge use but on understanding it so as to learn how to increase utilization success. Larsen (1980, p.425) discusses problems with direct observation. Dunn (1983) reviews various measurement procedures, including naturalistic observation (such as ethnography), content analysis and questionnaires and interviews.

[14] Rich (1975) and Caplan et al. (1975) distinguish conceptual use from instrumental use. Instrumental use has been defined as "knowledge for action"; it refers to instances where "respondents...could document the specific way in which...information was being used for decisionmaking or problem-solving purposes." By contrast, conceptual use has been defined as "knowledge for understanding"; it refers to instances where information "influenced a policymaker's thinking about an issue without putting information to any specific, documentable use" (Rich, 1977, p. 200). Whiteman (1985b) suggests that instrumental use should be called concrete use, to provide a more intuitive antithesis to conceptual use, while Dunn (1983) distinguishes conceptual and behavioral effects of knowledge.

shall now examine seven purposes and the meaning of utilization success for each. Note that the purposes correspond roughly to phases of the policymaking process.

1. The first purpose is *problem formulation.* Here utilization success occurs when participants in the policymaking process use study knowledge to shape or elaborate their conception of the policy problem: the objectives, the issues, the boundaries and constraints, the relevant consequences of policies and the ways to measure them, the important uncertainties, the relations among problem components, and so forth. (Section 4.3 and Chapter 5 of the *Overview* discuss problem formulation, but from the perspective of the analysts.)

While a systems analysis generally provides a problem formulation, participants in the policymaking process may not choose to use it; they may cling to a prior conception of the problem or may prefer one developed among themselves. On the other hand, they may derive part of their formulation from a systems analysis finding or some other aspect of study content that is not part of the analysis's formulation of the problem; for example, participants in policymaking might create a problem formulation that deliberately neglects a particular uncertainty because a prior systems analysis has shown through sensitivity analysis that the uncertainty had negligible effects.

A systems analysis may be judged successful if it influences the participants' problem formulation, individually or collectively. Usually this influence is partial, as when the problem statement or several other elements from the systems analysis are adopted. Consider three, quite different, examples from SANCAP. First, the SANCAP approach and terminology for describing alternatives hierarchically in terms of tactics, pure strategies and mixed strategies were used in the policymaking process to identify points of leverage. Second, the SANCAP analysis helped the policymakers pull in the problem boundaries to focus on strategies for controlling reactive hydrocarbons rather than other species of pollutants. The analysis demonstrated, for San Diego, the dominance of strategies that control reactive hydrocarbon emissions for controlling other species of emissions as well; that is, a strategy that

reduces reactive hydrocarbon emissions sufficiently to meet the oxidant standard will have already reduced all other species sufficiently to meet their respective standards. Third, the SANCAP analysis helped the policymakers to relax a major constraint in their formulation. The necessary reduction in pollutants to meet the oxidant air quality standard was determined originally on the basis of a certain historical oxidant concentration. When the analysis led to this concentration being rejected as anomalous by San Diego County and EPA Region IX, the necessary reduction became substantially smaller, making control strategies feasible that were less stringent and thus less costly (as much as $50 million per year less costly). (Goeller et al. 1973a, pp. 62,63)

An analysis's influence on problem formulation is not always partial, however. Sometimes, essentially, the complete problem formulation from a systems analysis is adopted in the policymaking process. This happened with POLANO and PAWN, and with the nominators and selectors--but not the EPA ratifiers--in SANCAP.

2. The second purpose is *generation of alternatives*. Here a systems analysis may be judged successful if it influences the nature or specification of alternatives considered by participants in the policymaking process. An analysis need not influence all the alternatives being considered, or all the participants considering them, to be deemed successful; it should usually be sufficient to influence a few alternatives, or components of alternatives, considered by a few participants.

This influence can happen in several ways. First, the participants may consider existing ideas for alternatives that the analysis team has systematically collected and presented. Examples include the technical and managerial tactics considered in PAWN and the retrofit emission-control devices for used vehicles considered in SANCAP. An unusual example comes from POLANO. About six months before the study began, the Cabinet and Parliament had declared, in principle, that they favored a storm-surge barrier--provided the Rijkswaterstaat could devise one that met certain conditions (such as acceptable levels of cost and flood security). Unless the conditions could be met, the mouth of the

Oosterschelde would supposedly be closed with an impermeable dam. Besides these two general alternatives, the analysis team investigated a third--leaving the mouth of the Oosterschelde open and constructing a system of large dikes around the perimeter of the estuary. They did this on their own initiative because, as Goemans (1986, p.7) notes,[15] they believed that "the knowledge base for decision-making should be as complete as possible". The analysts expected this alternative to be omitted from the Rijkswaterstaat's report to the Cabinet and Parliament because it went beyond the issue posed by Parliament. However, after seeing the three alternatives compared in a common framework, the Rijkswaterstaat boldly elected to include the open alternative in their report, where it was considered by both the Cabinet and the Parliament.

Second, the participants may consider new ideas for alternatives that were either invented by the analysis team or synthesized by them from existing ideas. For example, PAWN policymakers considered a new managerial policy for flushing the Markermeer to reduce its salinity, several new technical tactics for treating water supply problems, and a new managerial strategy for operating the network of major water management facilities, all of which were invented by the PAWN analysis team (Goeller et al. 1983, pp. 217, 231, 287). And SANCAP policymakers considered candidate retrofit strategies (promising combinations of emission-control devices for retrofit on used vehicles) and promising mixed strategies (that is, overall strategies) that were synthesized by the SANCAP analysis team.

Finally, the participants own creation of alternatives may be stimulated by study content. They may create alternatives that are hybrids or descendents of alternatives presented in the systems analysis. They may create alternatives that are intended to fill needs or holes identified by the analysis, as when PAWN found the previous Dutch national policy of phosphate control would be ineffective by itself in solving the eutrophication problem. And they may create

---

[15] Goemans, who was then with the Rijkswaterstaat, served as the POLANO project leader in the Netherlands (B.F. Goeller was the project leader at RAND), helped write the Rijkswaterstaat report, heard the Parliamentary debate, and later interviewed members about their reaction to the study.

alternatives that reflect systems analysis findings or other results, as when they concentrate on the kind of alternatives that the analysis has shown to be cost-effective or slight issues that the analysis has shown to be unimportant or intractable. For example, after the SANCAP analysis found that fixed-source controls were much more cost-effective than retrofit emission-control devices and both were much more cost-effective than transportation management measures, strategies devised for San Diego by EPA Region IX appeared to reflect this finding. As another example, PAWN found that "dilution was no solution to pollution"; that is, no national policy to redistribute the water would appreciably improve water quality nationwide, although there might be highly localized changes. This finding led participants in policymaking to slight such policies and to emphasize instead regional approaches to pollution problems.

3. The third purpose is *estimation of consequences.* Before they can properly compare and select among alternatives, participants in the policymaking process need relatively comprehensive estimates of the alternatives' potential consequences--consequences for the problem situation, for persons affected by the situation, and for other parties at interest.

A systems analysis may be judged successful for this purpose if (some) participants use its results or methodology to estimate alternatives' consequences. This need not happen for all, or nearly all, alternatives or consequences for an analysis to be deemed successful. Yet an analysis that supplies a large fraction of the estimates participants believe they need will probably be considered more successful than one that supplies a small fraction.

Participants may use an analysis in several ways for the purpose of estimation. The most obvious way is to adopt its estimates. POLANO, PAWN, and SANCAP provided most of the estimates used in the associated policymaking, although some came from complementary studies such as the Rijkswaterstaat's technical feasibility study of the storm-surge barrier. Goemans (1986, p. 8) notes that POLANO's systems analysis "report [Rijkswaterstaat, 1976] was actually used in both the Cabinet and Parliament; the [values of the] impacts [consequences] were hardly questioned and the discussion focused on preferences."

Another way is to use the methodology developed in the analysis to estimate the consequences of alternatives developed in the policymaking process. For example, the SANCAP methodology was used to estimate, in a common framework, the consequences of the EPA Region IX Final Implementation Plan (October 15, 1973) for the San Diego Air Quality Region.

Still another way is for participants to use findings or results to understand the issues or appreciate the influence of assumptions. For example, SANCAP sensitivity analysis showed that shifting the technical assumptions used to calculate emissions from the reference values considered realistic for San Diego to the more pessimistic values defined by EPA Region IX would greatly increase the cost and extent of strategies to meet the air quality standards. As an additional example, PAWN found that farmers, in the absence of government restrictions, might triple the number of sprinklers--and the corresponding demand for water--while seeking to maximize their profits. And, as a last example, POLANO's "disaggregate way of presentation using scorecards enhanced understanding of the issues and did much to get the idea accepted that there [was] no 'best' solution; the explicit treatment of uncertainties...and...sensitivity analysis (e.g., for various...[size openings in]...the barrier...) was useful to give a feeling for the relative influence of assumptions." (Goemans, 1986, p. 8-9) [16]

And the final way is for participants to recognize from the study the need for certain information that it lacks, and then to obtain and use this information. The information might be consequence estimates that the participants are better equipped to produce (such as the reactions of other actors), or information that must be obtained before particular consequences can be estimated or the corresponding estimation methodology completed. There is also a possibility that the information's potential usefulness might not have been appreciated by

---

[16] For examples of scorecards, see Tables 10.4 and 10.8 and the accompanying discussions on pages 385-388 and 413-416, respectively in *Craft Issues*. The *Overview* (pp. 96-99) also discusses scorecards and their rationale and presents examples from POLANO in Section 3.4.

the analysts; the participants may have recognized this only after examining the information the analysts did provide. [17]

4. The fourth purpose is rejection of alternatives. A systems analysis may be judged to be successful if it helps policymakers to reject inferior alternatives, particularly if they had be strongly supported by parties who expected to gain from their adoption. To someone not experienced in systems analysis, this purpose may seem less important than some of the others, but experience tells us that it is quite common in large organizations for courses of action to be strongly advocated--and even sometimes adopted--that will not have the desired effects, with a consequent waste of energy and resources. Thus, it is fairly important to help reject the really inferior alternatives, and much successful systems analysis work has done not much more than this.

Although the SANCAP and PAWN analyses did considerably more, they were quite successful at getting inferior alternatives rejected. Much of this success resulted from extensive screening analysis--described in *Craft Issues*, Sections 10.4 and 6.3--which identified alternatives that were either unpromising (not cost-effective) or dominated by other promising alternatives. PAWN's screening of technical and managerial tactics led to the rejection of most of the 57 national and regional tactics proposed, including several with construction costs over $100 million.

Rejection of inferior alternatives does not always result from screening, however. It also results from comparing alternatives' estimated consequences in detail. In SANCAP, for example, the San Diego County policymakers rejected alternatives D and E because their consequences were similar to but dominated by alternatives B and C respectively (see Table 10.8 in the *Overview*).

---

[17] The careful reader may have noticed that I focus on "policymakers" here, rather than on "participants in the policymaking process", as with previous purposes. What is important here is whether the analysis helps policymakers collectively decide to reject one or more alternatives, not whether the analysis influences a minority of policymakers or a majority of other participants (such as staff members or lobbyers) to argue for rejection.

5. The fifth purpose is *selection of a preferred alternative.* From the many alternatives considered in the policymaking process, one must be selected for eventual implementation.[18] A systems analysis may be judged a success here if it has a significant influence on the selection of the preferred alternative by policymakers.

Such success has several dimensions. The first dimension is the significance of the study's influence; that is, did it have an essential influence on the selection or merely an important one? This can be established by the answer to the question: Would the selection have been different without the study? If the answer is "probably so," then the study should be considered an essential influence on the selection. If the answer is "possibly so," then the study should be considered an important influence.

This dimension is relevant whether there is one policymaker or more. The other dimensions are relevant only when there are multiple policymakers, as with a committee, the U.S. Congress, the San Diego Board of Supervisors, and the Dutch Cabinet or Parliament.

The second dimension is the extent of the influence; that is, what proportion of the selections by individual policymakers did the study influence? A majority? A few? Of course the policymakers need not agree on their selections for the study's influence to be considered extensive; even if the study influenced all the policymakers, equal numbers might come to prefer different alternatives.

The final dimension is the decisiveness of the influence; that is, did the study influence the policymakers' collective selection? Although this may happen when the study influences a majority of

---

[18] Occasionally, where the implementability or the potential implementation costs of the preferred alternative remain a matter of concern after the analysis, two or three different alternatives may be selected provisionally and ranked by preference. If implementation planning subsequently showed unacceptably large implementation difficulties or costs for the alternative with highest ranking, then the next-highest ranked alternative that seems acceptable probably will be implemented (unless the implementation planning results appeared likely to change the rankings and the policymakers were willing and able to reconsider their preference).

policymakers to select the same alternative, it may also happen when the study influences only a few policymakers. The study might influence the few deciding votes in circumstances where most votes were divided equally between two alternatives for reasons other than the study. Or it might influence a few leaders among the policymakers to cause the group to agree on a selection where most members were motivated by the leaders rather than the study.

Multiple policymakers introduce an additional complication. Success in these dimensions may differ with the policymaker's subrole (nominator, selector, or ratifier).

The PAWN study's success was clear-cut. The study's influence on selection was essential, extensive, and decisive for policymakers in all subroles. The Minister of the Netherlands Ministry of Transport and Public Works, Mrs. N. Smit-Kroes (1984), stated that

> the new national water-management policy for the Netherlands is based largely on the PAWN project. Without the PAWN methodology and analysis, or something equivalent, to assess the cost and benefits of alternative policy actions in a credible way, many of the changes to the previous policy would not have been made.

As illustrations, consider several components of the national water management policy that were selected on the basis of PAWN (Goeller and the PAWN Team, 1985). First, although bids had gone out for constructing the Waddinxveen-Voorburg Canal, the Brielse Meer pipeline is being built instead, on PAWN's recommendation. This will yield a $38 million investment saving plus $15 million per year average net benefit in reduced salinity damage to agriculture.

Second, the new flushing policy in PAWN for the Markermeer has been implemented, yielding expected net benefits of between $1.2 and $5.4 million annually.

Third, there has been a drastic change in the Dutch approach to eutrophication, the next most serious water-quality problem. Because PAWN concluded that the previous Dutch national policy of phosphate control would be ineffective by itself, the Netherlands adopted a

regional approach to pollution control. Complementary tactics are to be added to the continuing national policy of phosphate control.

And finally, the national policy recommends that water board plans and the regional technical and managerial tactics identified as promising in PAWN's screening analysis be seriously considered during the development of provincial water-management plans. Implementation of all recommendations would produce expected profits of between $53 million and $128 million per year.

For SANCAP, success was almost as clear-cut. The study's influence on selection was essential, extensive, and decisive for the nominators and selectors; on the basis of the SANCAP analysis they chose an overall strategy (and technical assumptions for evaluating its emissions) developed in SANCAP in preference to those developed by EPA Region IX. For the ratifier, EPA Region IX, SANCAP had an important influence on part of the strategy and some of the assumptions selected; but it did not appear to have much affect on the ratifier's choice of transportation management controls, where there were some major differences from San Diego's selection.

For POLANO, the reception of the study was auspicious. The results were summarized by RAND in an all-day briefing to the Rijkswaterstaat on April 5, 1976, one year after the study began. Events then moved quickly to a conclusion. The Rijkswaterstaat (1976) report, based largely on POLANO, was presented to the Cabinet one month later, and then to Parliament, along with the Cabinet's selection of the storm-surge barrier alternative, which Parliament approved in June 1976.

But success was not as clear-cut for POLANO as for the other illustrative cases. On the one hand, the study had an essential, extensive, and decisive influence on the selection of a key storm-surge-barrier design component: the gates. The storm-surge barrier has sixty enormous gates that are normally open but can be closed to prevent large storm surges from entering the Oosterschelde. Each gate consists of a slot (the gate opening) that can be closed with either one or two slides. Whether single or double slides should be used was a crucial design question for the barrier. Double slides had been standard practice for other barriers in the Netherlands so as to minimize the

risk of flooding if a slide failed to close on command. But single slides were used in the storm-surge barrier, saving about $200 million, because POLANO showed they offered a very high level of safety and that it was doubtful whether double slides could increase this sufficiently to justify their much larger costs.

The study also had an essential, extensive, and decisive influence on the decision by the nominator--the Rijkswaterstaat--as to which alternative to recommend to the selector (the Cabinet) and how to pose the recommendation. The Rijkswaterstaat chose to mirror the POLANO study conclusions in its report (Rijkswaterstaat, 1976). By intention, the study did not conclude by recommending a particular alternative. Rather, it attempted to clarify the issues by comparing, in a common framework, the many different impacts of the alternatives, but left the choice among alternatives to the political process, where the responsibility properly resides. Significantly, there was no dominant alternative--one that was best for all the impacts. Indeed, each case had a major disadvantage that might be considered sufficiently serious as to render the alternative politically unacceptable: The storm-surge barrier case was by far the worst for cost; the closed case for ecology; and the open case for flood security.

On the other hand, it is hard to appraise the study's influence on the selection of the storm-surge-barrier alternative by the Cabinet and the Parliament (the ratifier), for several reasons. First, the Rijkswaterstaat's report to them did not recommend a particular alternative. Second, the privacy of the Cabinet and the size and complexity of Parliament handicap those who try to understand the process leading to a particular decision. It is clear that the POLANO results, embodied in the Rijkswaterstaat report, were an essential factor in Cabinet and Parliamentary deliberations. But it is not clear whether this was also true for their selections. Some observers believe it was, while others do not. Finally, since the Cabinet and Parliament favored the storm-surge barrier before POLANO began, provided certain conditions could be met, it is difficult to determine whether POLANO influenced the selection more than it helped confirm or justify a selection that had already been made on other grounds.

Even when it is high quality and potentially useful, a systems analysis may fail to have a significant influence on the selection of the preferred alternative. One reason is the political nature of policymaking. As Springer (1985, p. 490) points out, policymakers

> have obligations to honor *past commitments*, *to balance opposing values*, and to respond to the political give and take that is part of the institutional "cement" of our pluralist system. If specific analytic recommendations do not translate directly to action in this setting, it does not necessarily mean that the analysis "failed." It may simply have lost to some more pressing, or deserving, claim.

Another reason is that, compared to other sources, systems analysis may be at a disadvantage in supplying the type of information that turns out to be decisive, or especially important, in a particular selection process. Sabatier, quoted in Springer (1985, p. 491), contends that there are four other "types of information that a politically rational agency can be expected to pursue" besides the type of substantive information on policy problems and impacts typically generated by systems analysis. These competing types are (1) legal rules and regulations that affect agencies' decisions, (2) the history of agency decisions, (3) the "preferences of important actors," and (4) "the probable reactions" of these actors to the estimated "consequences of each policy alternative."

Sabatier's list has some communality with one by Brewer and deLeon (1983, pp. 190,191), who suggest that a policymaker "probably considers, either consciously or unconsciously, at least five factors before making a choice:" The context of the problem,[19] the points of leverage (that is, variables manipulable by policymakers), the availability (and trustworthiness) of information, the personalities of the participants,

[19] The context, as they define it, includes such questions as, "How is the problem defined and bounded?...When or how soon must the decisions be made? What are the rules of the game and the players?....[And] on what precedents is the decisions based, and how do they relate to present conditions?"

and the importance of the problem (which "governs how much of the decisionmakers's limited time, attention, and political resources will be expended on a given policy"). Again, systems analysis may be at a comparative disadvantage in supplying information on some of these factors.

6. The sixth purpose is *promotion of a preferred alternative.* A systems analysis may be judged a success for this purpose if information from the study is used to "advocate, [justify or] reaffirm policy positions after they have been determined."[20] It does not matter whether a position was influenced by the study or predetermined.[21]

For a study to be successful for this purpose, it must be used without distortion, which can occur through misrepresentation or selective use of arguments out of context. Distorted use is no credit to a study.

Consider some examples of promotion. Promotion occurs when an individual policymaker uses information from the study to advocate his preferred alternative to another policymaker.[22] And it may occur when he uses it to confirm or reinforce his original preference.

Promotion also occurs when policymakers collectively use information from the study to advocate or justify their position to another policymaking body. In SANCAP and POLANO, for example, the selectors used study results in advocating and justifying their preferred alternative to the ratifiers. In POLANO, after Parliament had

---

[20] Although Whiteman (1985a, p.200) calls this type of knowledge use "strategic," we feel it is better described as "promotion."

[21] Weiss (1977a, p.15) argues persuasively that using research "to support a predetermined position" is "neither an unimportant or an improper use...just because sides have already been taken is no reason to discount the effects of research." However, Pelz (1978, p. 351) warns of "symbolic" use, where performing a study is a substitute for a decision or where study knowledge is distorted to support publicly a decision that was predetermined or made on other grounds.

[22] If the other policymaker uses information conveyed by such advocacy in formulating the problem, estimating consequences, or rejecting or selecting alternatives, then information from the study is also being used for those purposes--and may contribute to corresponding successes.

ratified the Cabinet's selection of the storm-surge barrier, members of both groups used study results to justify their position to the general public and interest groups such as SOS.

7. The final purpose is *implementation*. A system analysis may be judged successful if it helps to implement the selected alternative.

This is almost automatic when the analysis has a significant influence on the alternative's selection and implementation ultimately occurs. Such was the case for POLANO, PAWN, and SANCAP. The Queen of the Netherlands formally dedicated the fully operational storm-surge barrier on October 4, 1986, in a ceremony attended by 25,000 people, including foreign dignitaries such as the Presidents of France and West Germany, and televised throughout the Netherlands. Nearly all the national water management policy developed through PAWN has been implemented, although some implementation is still in process. And, after multiple rounds of negotiation and revision, taking several years, a clean air strategy derived partly from SANCAP has been instituted in San Diego.

Even when it was not a significant influence on the alternative's selection, a system analysis may help implementation, and particularly implementation planning, in several ways. One way is through carryover of knowledge. An analysis generally explores the implementability of various alternatives, as well as their estimated consequences, and this knowledge may be useful in implementation.

Another way is by identifying issues for resolution in the final detailed design or plan for the selected alternative. Some POLANO examples include identifying the need for a storm-surge barrier control strategy and raising the possibility of dike redesign as an inexpensive means of hedging against uncertainty in certain assumptions.

Yet another way is by identifying data or research needed to facilitate implementation. For example, POLANO recommended additional research on scour processes (which potentially could undermine the barrier) so that barrier design modifications could be made, if necessary, with greater understanding. POLANO also recommended measuring the amount of nutrients being imported into the Oosterschelde by the North Sea. This quantity would affect the amount of biological

life the Oosterschelde could support with different-size openings in the barrier; also, if much of the import occurred during storms, it could restrict the choice of an operating policy for the barrier.

And the final way is by providing analysts knowledgeable about the problem who may participate in implementation planning or become implementation managers. For example, H.N.J. Smits, who was one of the Dutch analysts working on POLANO, subsequently worked on planning for the barrier and then became head of the Rijkswaterstaat department responsible for its planning and financial management. The analysis team's background makes it well suited to help constructively in devising ways of overcoming the difficulties that are bound to arise in an implementation.

Conclusions on Utilization Success. Utilization success is complex; a study may succeed in some purposes and fail in others, be used by some clients and not by others. And different users may use different elements of a study for different purposes. Utilization success, if it occurs, generally involves partial use of a study's content rather than complete use. Utilization success is thus a matter of degree.

The framework for utilization success offered here has important implications for those who would empirically measure knowledge utilization: It is difficult to measure the degree of success of one study. And it is more difficult to compare the success of several studies. The framework also suggests that empirical investigations of utilization must cast a wide net, lest they miss uses or users. Fortunately, the concern in this paper is not with measuring knowledge use but rather with understanding it sufficiently so as to learn how to increase utilization success. However, I believe that this framework can help those concerned with utilization measurement.

Outcome Success

While utilization success considers how the study was used in policymaking, outcome success considers what happened to the problem situation and those affected by it as a result of this use. Specifically, a systems analysis may be judged an outcome success if two conditions are satisfied: (1) the implemented alternative improves the

problem situation, and (2) the collective selection of that alternative was significantly influenced by the analysis.

The second condition is identical with utilization success for the purpose of selection, which was discussed above. The first condition is the subject of evaluation research. Since Chapter 11 of *Craft Issues* offers a useful overview, there is no need to summarize the burgeoning evaluation-research literature here (for two excellent but different overviews, see Brewer and deLeon, 1983, part V, and Patton and Sawicki, 1986, Chapter 9). It is useful, however, to discuss briefly several evaluation-research complications that make it very difficult to determine whether the first condition has been satisfied--and thus whether outcome success has been achieved.

The first complication is attribution of cause. Simple comparisons of the problem situation before and after policy implementation often cannot establish whether the policy actually caused any or all of the observed differences; the differences might result instead from changes in the environment or other factors. This complication has sometimes fostered experimental-design approaches to evaluation where the policy is applied to one situation or group and withheld from another (the control). Unfortunately, such experimental approaches are impractical for many policy problems.

The second complication is the measurement of multidimensional outcomes. The outcome of an implemented policy generally involves many and varied consequences, which differ in their ease, expense, and accuracy of measurement. Resource limitations may force unfortunate compromises in outcome evaluation. If an evaluation measures only a few potential consequences, it is limited and may miss important shortcomings of the policy. Suppose, for example, that an evaluation of SANCAP concentrated on the policy's performance in improving air quality. It might miss the fact that the implemented policy led to a subtle but serious degradation in transportation service for a particular segment of the population. On the other hand, if an evaluation tries to measure all the potential consequences of a policy, the evaluation costs would probably be overwhelming. And if one tries to economize on measurement processes, the measurements may be so inaccurate as to be misleading about a policy's actual outcome.

The third complication is balancing multiple evaluation criteria. "Considering the number and diversity of possible outcomes and effects for even the simplest program, one nearly always must use multiple criteria to evaluate" (Brewer and deLeon, 1983, p. 331). But the existence of multiple evaluation criteria leads to the problem of how to balance or weigh diverse and often conflicting criteria. (This is conceptually similar to the multicriteria problem that policymakers face when selecting among alternatives, as discussed in *Craft Issues*, Section 8.4). The problem is aggravated by the fact that the implemented policy may have both favorable and unfavorable consequences, and some consequences may be larger and others smaller than expected or desired.

The fourth complication is the standard of comparison. Evaluations generally have a standard of comparison, explicit or implicit, that can affect the conclusions. Many evaluations compare actual outcomes to expected outcomes (Brewer and deLeon, 1983, p. 20), which often reflect the analysis estimates on which the policy was selected. This standard primarily evaluates the accuracy of the expectations--and the underlying estimates.[23]

Problem situations evolve with time, of course, so that by the time a systems analysis is complete and a resulting implementation is in place, the situation may have changed enough for the outcome to be somewhat different from that predicted by the analysis. In such circumstances, the analysis team can sometimes adjust the analysis results for the changed situation so the comparison with the actual outcome will be a fair one.

A different standard is whether the affected people consider the actual outcome attractive (desirable or even acceptable). This attractiveness standard is developed after the policy has been implemented and taken effect (that is, during evaluation), while the evaluations standard is developed before implementation. Thus, the

---

[23] This comment usually applies to a related standard that compares actual outcomes to policy objectives. The policy objectives for the selected policy are usually established on the basis of systems analysis estimates. Sometimes, however, they are general objectives that were established before the analysis.

expectation standard can evaluate the attractiveness of a policy's outcome to the affected people only to the extent that the policymakers considered the people's preferences (preference functions) during selection and these preferences have remained unchanged. The longer the period between policy selection and evaluation, the greater the possibility that these preferences will change.

The fifth complication is the multiplicity of audiences. Evaluations have multiple audiences with different concerns. Policymakers are concerned with the implemented policy's performance in the problem situation. Is it effective or should it be terminated because of ineffectiveness or unacceptable side effects? Implementation managers are concerned with where the policy is having problems. Where and how can policy implementation and operation be improved? Interest groups of affected persons are concerned with how the policy is affecting their group in comparison to others. How are the actual consequences distributed among different groups? Is my group getting its fair share or bearing an unfair burden? (Interest groups are much more interested in policies' distributional effects than other audiences.)

Given limited resources, each evaluation must limit itself as to intended audience and included concerns. Depending on this choice, an evaluation may not be helpful for judging outcome success. Evaluations focused on implementation managers are a prime example. Such evaluations emphasize the issue of how to improve the policy rather than the issue of how the policy, as originally implemented or modified, has improved the problem situation.

The sixth complication is the timing of the evaluation. If an evaluation is performed too soon after implementation, there may not have been sufficient time for the policy to take effect and its resultant consequences to become apparent. For POLANO, as an example, the Rijkswaterstaat(1986) performed an evaluation 10 years after the selection of the storm-surge barrier alternative. Although implementation had proceeded with all deliberate speed, the barrier was still a few months short of full operation. The Rijkswaterstaat evaluation report (1986) evaluated "the extent to which the predictions

... conform[ed] to reality" by comparing the observed outcomes with the predictions in the "White Note", the Rijkswaterstaat report (1976) to the Cabinet and Parliament that was largely based on POLANO.

The evaluation report found that observations showed (essentially) no difference from predictions for most consequences, improvements for several consequences, and declines for two consequences: financial cost and completion date. The cost estimates for the overall project were overshot by 10 percent, and those for the storm-surge barrier by 30 percent. (The completion date was overshot by a year, partly because of a political decision to delay completion to reduce budgetary impacts.)

The report skipped the evaluation of two important aspects of the problem--the marine ecology and the fishing industry--because the policy's consequences for these aspects would not be apparent until perhaps a half dozen years after completion of the construction.

And while the report attempted to evaluate flood security, it lacked the necessary data for a complete job. During the first fifteen months of its operation, according to Dye (1988), the barrier "has closed off the sea at least three times". But the security of the barrier has not been severely tested. The barrier, after all, was designed to provide protection against a storm so severe that it might be expected to occur only once in 4000 years. Flood security thus provides an extreme illustration of how some consequences cannot be fully evaluated within any reasonable time horizon.

Of course an evaluation can also be performed too long after implementation. Long delays complicate attribution of cause, and make some improvements in consequences unmeasurable. And, as mentioned earlier, the longer the period between policy selection and evaluation, the greater the possibility that the preference functions of policymakers or interest groups may change.

The final complication is the cost of evaluation. Good evaluation is costly. Many implemented policies will not be deemed important enough to warrant the resources--time, money, and talented researchers--that a detailed evaluation demands. And, as Brewer and deLeon (1983, p. 360) point out, evaluations may have unacceptable political costs; they "disrupt, reveal institutional weaknesses and limitations, and threaten individual policies and their sponsors."

Because of the complications we have discussed, an implemented policy will often receive no evaluation, or an evaluation that cannot establish whether or not it improved the problem situation. Thus, it is difficult for a systems analysis to satisfy the first condition for outcome success mentioned above. When this difficulty is compounded with the difficulty of satisfying the second condition, the clear conclusion is that it is extremely difficult for a study to be judged an outcome success--even when it truly is one.

On the other hand, if a policy clearly fails to achieve its goals or improve the situation, it does not necessarily mean that the reason for this outcome failure was either a defective analysis (theory or assumptions) or a defective selection. Among other reasons, it may be, as Schneider (1986, p.357) suggests, because "the agencies responsible for implementation failed to properly translate the policy theory into an operational plan or [because] inadequate resources were available to those responsible for implementation."

Indirect Success

Thus far this section has discussed success on a direct basis, where one assesses the three kinds of success with respect to the decision and problem situation for which the analysis was commissioned. It addressed such questions as: How were the POLANO methodology and results used in selecting a flood protection alternative for the Oosterschelde? And how were the SANCAP methodology and results used in selecting a strategy to clean up the air in San Diego?.

Now the paper will discuss success on an indirect basis, that is where one accesses success with respect to *other* decisions or situations. The discussion will address such questions as: How were the POLANO methodology and results used in selecting a strategy to control the storm-surge barrier in the Oosterschelde? And how were the SANCAP methodology and results used in selecting a strategy to clean up the air or improve transportation in Los Angeles? Because this section has previously considered the three kinds of success in detail, the discussion of them on an indirect basis will be fairly brief, mainly providing examples and describing unusual features.

Utilization Success. For utilization success, a study must be used in the policymaking process. When study methodology or results are used for some purpose in the policymaking process of another decision or problem situation, it counts toward indirect success for that purpose.[24] This is easily clarified with examples.

The POLANO methodology was used for several purposes in other decisions and problem situations. A refined version of POLANO's SIMPLIC model for estimating water levels was used in the BARCON project to generate alternative control strategies for the storm-surge barrier and estimate their consequences. It was also used by civil engineers developing the detailed design for the barrier.

The POLANO model of eutrophication (algae blooms) in salt water was extensively modified for PAWN, where it became the primary tool for developing and estimating the consequences of fresh-water eutrophication control strategies.

All the models in the SANCAP methodology, except the air quality model and the economic input/output model, were recalibrated and used for Los Angeles. (Some were extended to account for new problem features.) They were applied to generate alternatives and estimate their consequences in two major studies of Los Angeles; one study was concerned with alternative strategies for oxidant control (Goeller, et al. 1973b) and the other was concerned with near-term transportation alternatives for simultaneously treating air pollution, energy conservation, and transportation service problems (Mikolowsky et al. 1974). Both studies were utilization successes. The oxidant-control study led to several inferior alternatives being rejected; the transportation study produced a strategy (strategy E) that became the basis of the short range transportation plan adopted for Los Angeles (Southern California Association of Governments, 1974).

But methodology was not the only SANCAP product carried over to these Los Angeles studies. In constructing alternatives, they used the fixed source and retrofit strategies developed in SANCAP, along with the

---

[24] Collateral success is another term for indirect success.

SANCAP finding about the relative cost-effectiveness of these strategies and transportation management measures in controlling emissions.

Much of the SANCAP problem formulation (for example, issues, relevant consequences of policies and how to measure them, important uncertainties, relations among problem components) also carried over to the Los Angeles studies, although some region- and problem-specific items did not. The SANCAP approach and terminology for describing alternatives hierarchically in terms of tactics, pure strategies, and mixed strategies was not only used in the Los Angeles studies, but also in PAWN. In this way, and others, PAWN also counts toward SANCAP's indirect success in problem formulation.

PAWN identified a number of water board plans and regional technical and managerial tactics as promising during screening (see *Overview*, Section 6.3), all of which were recommended by the national water management policy. Their implementation, which is beyond the authority of national policymakers, is being seriously considered during the development of provincial water management plans.

For many lakes in the Netherlands, the provinces are also using improved versions of the PAWN eutrophication models to develop a tailored combination of eutrophication control measures.

The comprehensive methodology PAWN developed has been adopted for national water resources planning by the Rijkswaterstaat and by the Delft Hydraulics Laboratory and has been used in several major studies. Parts of the methodology are being used by the Ministry of Agriculture and Fisheries, and by six of the 11 Dutch provinces.

As the previous example suggests, a study's methodology or results can produce "ripples" of use that spread more and more widely among decisions and problem situations.[25] This may happen haphazardly, by word

---

[25] For a particular study result, Dunn (1983) characterizes the scope of knowledge use as either specific (for example, the use of a study recommendation for a particular decision) or general (for example, ideas in wide use). Weiss, as cited by Larsen (1980, p. 425), employs the term "enlightenment" to describe the broad influence of knowledge, where "it is not the findings of a single study nor even a body of related studies that directly affect policy [but]...the concepts and theoretical perspectives that...research has engendered..." (see also Weiss, 1977b and 1980).

of mouth or by migration of the analysts or policymakers; or it may happen more systematically as the methodology and results become part of the state of the art or are propagated through the literature or courses. Tracing such ripples is difficult, and often impossible. Nevertheless, I can offer a few examples.

The use of findings is probably hardest to trace. This is partly because they often arrive at a new situation as part of general knowledge, so their source is unclear, and partly because they are often adapted to fit the new situation. For example, I believe, but cannot prove, that several SANCAP findings were used in developing other regions' strategies to clean up their air. These include: (1) the finding that fixed source controls were generally most cost-effective, (2) the finding that tractor-towing of jet aircraft within airports and the use of adsorption devices and substitution of nonreactive coatings in surface coating (such as painting) operations were promising measures to reduce emissions, and (3) the finding that some technical assumptions promulgated by EPA were unrealistic for certain regions and others were flawed. I also believe that EPA modified, or permitted the modification of, several assumptions on the basis of SANCAP and the Los Angeles oxidant control study findings.

Although the scorecard technique (*Overview*, pp. 96-99) was first used in another study in 1971, it was refined and propagated through SANCAP. It has now been used in a wide variety of problems throughout the world.

In SANCAP, screening received an early, and perhaps the earliest, application in a major systems analysis. Here it also became the first stage of a multi-stage process to generate and evaluate alternatives. (This process performs what has previously been referred to as "explicit hierarchical design and evaluation of alternatives;" see the discussion in *Craft Issues*, Chapter 6.) I believe that much of the subsequent use of screening and this multi-stage process in other studies stems from SANCAP and its progeny.

According to Goemans (1986, p. 12-13), the POLANO approach "attracted much attention. Somewhat reluctantly at first, but more enthusiastically in later years, the scorecard was included in courses

for government officials. [Furthermore,] after the Oosterschelde controversy, the Ministry of Transport and Public Works and other Ministries felt the need for an independent institute...to do policy studies...." This led to the establishment in 1981 of the SIBAS Joint Institute for Policy Analysis, with Goemans, the POLANO project leader in the Netherlands, chosen to be its Director.

The Delft Hydraulics Laboratory has applied PAWN's general approach in master planning studies for Taiwan and several other countries, and in studies of shore protection for the Italian coast near Ravenna and the Dutch island of Texel. SIBAS has applied the approach in several studies, including an ongoing study of strategies to manage the Netherlands' section of the North Sea and its associated resources.

PAWN provides a thoroughly documented case study to educate decisionmakers and train analysts in the analysis of complex natural resource and environmental questions. It has already been used to train water resource planners in the Rijkswaterstaat and all 11 Dutch provinces, as well as planners from Bangladesh in a three-month course sponsored by the World Bank. It was also the basis of a two-week training course given in November 1984 for member nations of the United Nation's Economic and Social Commission for Asia and the Pacific. With changes to the data base, many models are transferable to other countries.

The illustrative case studies we have employed in this paper have appeared in various policy analysis books; for example, POLANO has appeared in the *Overview* and in Quade (1982); SANCAP has appeared in Mood (1983), in Quade (1975), and in Chapter 10 of *Craft Issues*; and PAWN appears in Chapter 6 of *Craft Issues* and will be discussed in the planned *Cases* volume of the Handbook of Systems Analysis.

Outcome Success. It has already been noted that it is extremely difficult for a study to be judged an outcome success on a direct basis--even when it truly is one. This difficulty is enormously greater on an indirect basis.

There are two general ways for a study that was commissioned for a particular decision and problem situation to become an outcome success on an indirect basis. One way is for one or more items (such as

findings, estimates, pieces of problem formulation or methodology) from the study to significantly influence the policymaking for another decision or problem situation, where that policymaking results in outcome success (as defined earlier). The other way is for one or more items from the study to sufficiently influence the results of a subsequent study for another decision or problem situation so as to be an important factor in the subsequent study's outcome success (on a direct basis).

I have no examples of the first way arising out of our illustrative case studies, partly because I have not tried to trace their influence carefully and partly because knowledge propagation can be diffuse and tenuous.

The best example of the second way is the study of near-term transportation alternatives for Los Angeles, mentioned above. Viewed on a direct basis, that study is an outcome success; the implementation of a strategy developed and selected largely on the basis of the study has apparently improved the problem situation. Furthermore, SANCAP strongly influenced the approach and results of the study. Most of the problem formulation and methodology carried over. And in constructing alternatives, the Los Angeles study used fixed-source and retrofit strategies developed in SANCAP, along with the SANCAP finding about their relative cost-effectiveness in comparison to transportation measures for controlling emissions. Viewed on an indirect basis, it therefore appears quite reasonable to count the Los Angeles study as an outcome success for SANCAP.

Another, less definitive example of the second way is the BARCON study. POLANO's SIMPLIC model for estimating water levels was the tool BARCON used to develop and estimate the consequences of control strategies for the storm-surge barrier, including the strategies that were subsequently selected and implemented. These strategies appear to work well and meet their goals, which means that BARCON would probably be considered an outcome success when viewed on a direct basis, although evaluation has been limited by the fact that the barrier began operation only about a year and a half ago. Viewed on an indirect basis, it therefore appears plausible to count the BARCON study as an outcome

success for POLANO. By contrast, it is not clear whether POLANO itself was an outcome success when viewed on a direct basis; this is because of the previously discussed uncertainty as to the extent POLANO influenced the selection of the storm-surge barrier alternative by the Cabinet and Parliament

Analytic Success. Recall that analytic success is concerned with the quality of a study's methodology and results. It is also the foundation for the other kinds of success. The utilization of poor-quality results or the implementation of a policy whose selection was based on poor-quality results is no credit to a study.

On a direct basis one assesses a study's quality of methodology and results with respect to the decision and problem situation for which the analysis was commissioned. But on an indirect basis one assesses the quality of methodology and results with respect to their use in *another* decision or problem situation. That is, one asks whether such use for a particular purpose is invalid or inappropriate. For example, is this SANCAP finding invalid for the Los Angeles near term transportation study? Is this POLANO estimate invalid or inappropriate for BARCON? Is the POLANO model inappropriate for PAWN? Is this SANCAP model, suitably modified and recalibrated, invalid or inappropriate for the Los Angeles oxidant control study? Findings and estimates have underlying assumptions that may be violated if they are translated to different circumstances. Models have ranges of validity that may be exceeded and calibration factors that should be superseded if they are applied to sufficiently different circumstances.

I believe that the studies mentioned above as utilization and outcome successes on an indirect basis are also analytic successes on an indirect basis. When a subsequent study borrowed from an earlier study, considerable care was devoted to insuring that what was borrowed was not invalid or inappropriate for the new circumstances. For example, models were generally recalibrated and revalidated. And findings were only translated to circumstances where the underlying assumptions were fairly well satisfied. Furthermore, the analysts on the earlier study often tried to anticipate other applications or circumstances where estimates, findings, or models might be applied and then to suggest which might be inappropriate.

While it is important to assess analytic success on an indirect basis to avoid misapplying a study's products, it is also important to learn how to apply the elements of one study to other decisions or problem situations whenever possible, so as to avoid reinventing the wheel.

CONCLUSIONS ABOUT SUCCESS

In the light of the framework for describing success set forth in the preceding section, it is appropriate to close this paper with some general conclusions about success in systems analysis.

In managing an analysis to achieve success, analysts should give primary emphasis to analytic success and secondary emphasis to utilization success viewed on a direct basis, largely ignoring outcome success and slighting success viewed on an indirect basis. This recommendation reflects conclusions concerning the relative importance and the analysts degree of control of each. Below there are several reasons for these conclusions.

Analytic success is the foundation of other kinds of success, and also the one over which analysts have the most control.

Measures of utilization and outcome success for a policy study necessarily confound the effects of study quality with the effects of political incentives and constraints. Even a study of "perfect" quality may not succeed in influencing policy, or even in being used, because some results conflict with political constraints or incentives. To the extent that such conflicts can be anticipated, the "perfect" study attempts to minimize them; it may try to undermine or remove constraints, and it may try to weaken conflicting incentives or to bolster supportive ones. Inevitably, some study findings will have conflicts that either could not be anticipated or about which the analysts can do nothing. Thus, a "perfect" study can only maximize its own potential; it cannot assure its own utilization (Fisher, 1987).

Nagel (1984) nicely supports the recommendation about the relative emphasis to give analytic success and utilization success: "The primary obligation of a policy researcher is to do valid research, rather than

research that is utilized. Greater sensitivity to the factors that facilitate research utilization can, however, be helpful in suggesting legitimate ways to increase the probability that valid research will be utilized." The wisdom of this view becomes clearer when one considers the implications if analysts viewed their primary obligation as utilization. Analysts would tend to tolerate or condone distorted use of their study results, either through misrepresentation or selective use of arguments out of context. And they would be tempted to perform analyses that pandered to policymakers' prejudices or *a priori* preferences among alternatives rather than analyses that sought the truth.

Outcome success is highly problematic. It is hard to establish whether the alternative's selection was significantly influenced by the analysis. Many things outside the analysts control can go wrong in translating the policy alternative into an operational plan and in implementing it. And evaluation's many complications make it extremely difficult to establish that a study is an outcome success even when it truly is one.

Finally, it is impossible to foresee most of the other decisions or problem situations where a study's methodology or results might be applied. Thus a study should concentrate on the decision and problem situation for which it was commissioned and seek direct success there; of course, in developing a study's methodology and results, it is worthwhile to give some attention to identifying other decisions and problem situations where the approach and findings might be used and to suggesting how they might be generalized, in order to facilitate indirect success.

One can explain the recommendation for giving primary emphasis to analytic success and secondary emphasis to utilization success by drawing an analogy to the saying: "The proof of the pudding is in the eating." In performing analysis we are primarily concerned with successfully cooking and presenting the pudding (analytic success)--how to equip and staff a kitchen, how to choose prime ingredients, and how to prepare, bake, and serve the pudding. If the pudding is not tastily prepared and attractively presented, it probably will not, and

should not, be eaten. We are secondarily concerned with successfully getting customers to eat the pudding (utilization success)--how to establish a trustful relationship with the customer, have a congenial serving staff, choose a popular flavor for the pudding, and ensure that the pudding complements the rest of the menu. If the pudding is not eaten, it can neither please the customer nor nourish him. We are not really concerned with suggestions for successfully using the pudding to improve the customers' nourishment (outcome success); too many factors besides pudding affect his nourishment for us to evaluate it properly. Nor are we particularly concerned with successfully enhancing what we or the customers do with leftover pudding (indirect success).

ACKNOWLEDGEMENTS

Rick Eden made important substantive and organizational suggestions. Kip Fisher provided valuable research assistance and stimulating dialogue. Steve Drezner gave essential intellectual and financial support. Gene Fisher carefully reviewed a preliminary draft. Hugh Miser offered helpful editorial suggestions as well as a thoughtful critique. Finally, Ed Quade, mentor and friend, provided constructive criticism and encouragement. It is to his memory that this paper is fondly dedicated.

REFERENCES

Armstrong, J. S. (1982). "The ombudsman: Is review by peers as fair as it appears?" *Interfaces* 12, pp. 62-74.

Beneviste, G. (1977). *The Politics of Expertise.* San Francisco, California: Boyd Fraser.

Berman, P. and M. W. McLaughlin (1977). *Federal Programs Supporting Educational Change, Vol. VII: Factors Affecting Implementation and Continuation*, R-1589/7-HEW. Santa Monica, California: The RAND Corporation.

Beyer, J. M. and H. M. Trice (1982). "The utilization process: a conceptual framework and synthesis of empirical findings." *Administrative Science Quarterly.* 27(4), pp. 591-622.

Brewer, G. D. and P. deLeon (1983). *The Foundations of Policy Analysis.* Homewood, Illinois: The Dorsey Press.

Caplan, N., A. Morrison and R. J. Stambaugh (1975). *The Use of Social Science Knowledge in Policy Decisions at the National Level.* Ann Arbor, Michigan: Institute for Social Research, University of Michigan.

Cases. See Miser and Quade (forthcoming).

Craft Issues. See Miser and Quade (1988).

Catlett, L., et al. (1979). *Controlling the Oosterschelde Storm-Surge Barrier--A Policy Analysis of Alternative Strategies: Vol. I, Summary Report,* R-2444/1-NETH. Santa Monica, California: The RAND Corporation.

Dunn, W. N. (1983). "Measuring knowledge use." *Knowledge.* 5(1), pp. 120-133.

Dye, L., "Barrier against the sea." *Los Angeles Times,* January 25, 1988, part II, p. 4.

Fisher, K. T. (1987). Personal Communication. Santa Monica, California.

Goeller, B. F., et al. *Protecting an Estuary from Floods--A Policy Analysis of the Oosterschelde: Vol. 1, Summary Report,* R-2121/1-NETH. Santa Monica, California: The RAND Corporation.

Goeller, B. F., et al. (1973a). *San Diego Clean Air Project: Summary Report,* R-1362-SD. Santa Monica, California: The RAND Corporation.

Goeller, B. F., et al. (1973b). *Strategy Alternatives for Oxidant Control in the Los Angeles Region,* R-1368-EPA, Santa Monica, California: The RAND Corporation.

Goeller, B. F., et al. (1983). *Policy Analysis of Water Management for the Netherlands: Vol. I, Summary Report,* R-2500/1-NETH, Santa Monica, California: The RAND Corporation.

Goeller, B. F., and the PAWN Team (1985). "Planning the Netherlands water resources." *Interfaces* 15(1), 3-33.

Goemans, T. (1986). *The Impact of the POLANO Study,* M-86165. Delft, The Netherlands: SIBAS.

Goldhamer, H. (1978). *The Advisor.* New York: North Holland.

House, P. W. (1982). *The Art of Public Policy Analysis.* Beverly Hills, California: Sage Publications.

Larsen, J. K. (1980). "Knowledge utilization: What is it?" *Knowledge.* 3(1), pp. 421-442.

Larsen, J. K. (1985). "Effect of time on knowledge utilization." *Knowledge.* 7(2), pp. 143-159.

Leemans, A. (1986). "Information as a factor of power and influence: Policymaking on the Delta/Oosterschelde sea defense works." *Knowledge* 8(1), pp. 39-58.

Lynn, L. E., Jr. (1978). "The question of relevance." In *Knowledge and Policy: The Uncertain Connection.* (Lawrence Lynn, Jr., ed.). Washington, DC: National Academy of Sciences. pp. 12-22.

Mikolowsky, W. T. et al. (1974). *The Regional Impacts of Near-Term Transportation Alternatives: A Case Study of Los Angeles,* R-1524-SCAG, Santa Monica, California: The RAND Corporation.

Miser, H. J., and E. S. Quade (1985). *Handbook of Systems Analysis: Overview of Uses, Procedures, Applications, and Practice.* New York: North-Holland. Cited as *Overview.*

Miser, H. J., and E. S. Quade (forthcoming). *Handbook of Systems Analysis: Cases.* New York: North-Holland. Cited as *Cases.*

Miser, H. J., and E. S. Quade (1988). *Handbook of Systems Analysis: Craft Issues and Procedural Choice.* New York: North Holland. Cited as *Craft Issues.*

Mood, A. M. (1983). Chapter 20, The RAND San Diego Clean Air Study. In *Introduction to Policy Analysis.* New York: North Holland.

Nagel, S. S. (1984). "MS/OR utilization in public policy making." *Interfaces* 14(4), pp. 69-75.

Overview. See Miser and Quade (1985).

Patton, C. V., and D. S. Sawicki (1986). *Basic Methods of Policy Analysis and Planning.* Englewood Cliffs, New Jersey: Prentice-Hall.

Pelz, D. C. (1978). "Some expanded perspectives on use of social science in public policy." In *Major Social Issues: A Multidisciplinary View.* (M. Yinger and S. Cutler, eds.). New York: Free Press.

Petruschell, R. L., et al. (1982). *Policy Analysis of Water Management for the Netherlands: Vol. XIII, Models for Sprinkler Irrigation System Design, Cost, and Operation.* N-1500/13-NETH. Santa Monica, California: The RAND Corporation.

Pulles, J. W. (1985). *Beleidsanalyse voor de Waterhuishouding in Nederland/PAWN.* The Hague, Netherlands: Rijkswaterstaat.

Pulles, J. W. (1987). Personal communication. Gouda, The Netherlands.

Quade, E. S. (1975). *Analysis for Public Decisions.* New York: American Elsevier.

Quade, E. S. (1982). *Analysis for Public Decisions.* 2nd Edition. New York: North Holland.

Ravetz, J. R. (1971). *Scientific Knowledge and Its Social Problems.* Oxford, England: Oxford University Press.

Rich, R. F. (1975). "Selective utilization of social-science related information by federal policymakers." *Inquiry.* 13(3). pp. 72-81.

Rich, R. F. (1977). "Uses of social science information by federal bureaucrats: Knowledge for action versus knowledge for understanding." In *Using Social Research in Public Policymaking.* (C. A. Weiss, ed.). Lexington, MA: D.C. Heath, pp. 197-212.

Rich, R. F. (1981). *Social Science Information and Public Policymaking.* San Francisco, California: Jossey-Bass.

Rijkswaterstaat (1976). *Analysis of Oosterschelde Alternatives.* (An English translation is available: *Policy Analysis of Eastern Scheldt Alternatives,* Nota DDRF-77.158. The Hague, Netherlands: Rijkswaterstaat.)

Rijkswaterstaat (1985). *De Waterhuishouding van Nederland.* The Hague, Netherlands: Rijkswaterstaat.

Rijkswaterstaat (1986). "En evaluatie van de witte nota." *Oosterscheldewerken/Voortgangsrapportage* 20. (January-June 1986.) (An english translation of this article entitled "An evaluation of the White Note" was prepared by M. Jas.)

Schneider, A. L. (1986). "The evolution of a policy orientation for evaluation research: A guide to practice." *Public Administration Review*, July/August 1986, pp. 356-361.

Shubik, M. (1984). "Gaming: A state-of-the-art survey." In *Operational Gaming: An International Approach* (I. Stahl, ed.). Oxford, England: Pergamon, pp. 15-22.

Smit-Kroes, N. (1984). Letter from the Minister of the Netherlands Ministry of Transport and Public Works. The Hague, Netherlands.

Southern California Association of Governments (1974). *SCAG Short Range Transportation Plan.* Los Angeles, California.

Springer, J. F. (1985). "Policy analysis and organizational decisions: Toward a conceptual revision." *Administration & Society.* 16(4), pp. 475-508.

Weiss, C. H. (1977a). Introduction. In *Using Social Research in Public Policy Making.* (C. H. Weiss, Ed.). Lexington, Massachusetts: D.C. Health. pp. 1-22.

Weiss, C. H. (1977b). "Research for policy's sake: The enlightenment function of social science research." *Policy Analysis.* 3(4), pp. 531-546.

Weiss, C. H. (1980). "Knowledge creep and decision accretion." *Knowledge.* 3(1). pp. 381-404.

Weiss, C. H. and M. J. Bucuvales (1980). *Social Science Research and Decision-Making.* New York: Columbia University Press.

Whiteman, D. (1985a). "The fate of policy analysis in congressional decision-making: Three types of use in committees." *Western Political Quarterly.* 38(2), pp. 294-311.

Whiteman, D. (1985b). "Reaffirming the importance of strategic use: A two-dimensional perspective on policy analysis in Congress." *Knowledge* 6(3), pp. 203-224.